SERGE LANG

# INTRODUCTION TO TRANSCENDENTAL NUMBERS

# INTRODUCTION TO TRANSCENDENTAL NUMBERS

BY SERGE LANG

*Columbia University*

The theory of transcendental numbers consists in determining the transcendence and algebraic independence of numbers obtained as values of classical functions, suitably normalized. This advanced text examines all of the several variations of the one main method of this theory. Applications range from a very elementary setting (concerning the function $e^t$), to rather sophisticated contexts involving abelian functions and automorphic functions.

More elementary sections are kept separate from the others, and therefore more accessible for the reader of limited background. When used as a text for a one-term course at the graduate level, certain portions of the theory should be omitted. Other parts can be covered according to the reader's capacity. For example, Chapter VI on transcendence measures involves more complicated techniques essential to further work in that direction.

90/-
£3.50)

# INTRODUCTION TO TRANSCENDENTAL NUMBERS

**SERGE LANG**
Columbia University, New York, New York

**ADDISON-WESLEY PUBLISHING COMPANY**
Reading, Massachusetts · Palo Alto · London · Don Mills, Ontario

This book is in the

ADDISON-WESLEY SERIES IN MATHEMATICS

LYNN H. LOOMIS

*Consulting Editor*


Printed in the United States of America.
Published simultaneously in Canada.
Library of Congress Catalog Card Number 66-28459.

Partially supported by NSF Grant GP5–303

# Foreword

The theory of transcendental numbers is reaching a stage where it is ready to take its place as one of the most attractive branches of mathematics. It consists in determining the transcendence and algebraic independence of numbers obtained as values of classical functions, suitably normalized. (We shall make this more precise in its appropriate place in the book.)

There is one main method, with several variations, which consists of constructing a function with many zeros out of the functions whose values one is considering. By *many* zeros one may mean just one zero with high multiplicity, or many distinct zeros with no condition on the multiplicity, or many distinct zeros with high multiplicities. We shall see examples of all three cases.

The theory has applications ranging from a very elementary setting (concerning the function $e^t$), to rather sophisticated contexts, having to do with abelian functions and automorphic functions. The embedding of "elementary" results in broad coherent theories has been known to cause acute cases of paranoia to persons who prefer "simple" significant examples to the elaboration of the more extensive theories. I personally like both, but I have made an attempt to keep the more elementary portions separate from the others, and hence more accessible for the reader of limited mathematical background, by:

1. Treating separately the special case of the exponential function, which needs only standard facts from basic courses in complex variables, and elementary properties of algebraic numbers.

2. Summarizing at the beginning of the book, with proofs, the few facts about algebraic numbers which we shall need in most of what follows.

Historical notes at the end of each chapter serve as much to describe past work in the subject as to suggest further possibilities and conjectures.

The book can be used as a text for a one term course in the theory of transcendental numbers, at the graduate level, if only certain portions of the theory are covered, e.g. Chapter I, Chapter II (omitting §4), Chapter III (omitting §4), the beginning of Chapter V, and Chapter VII. The other

parts of the book can be covered if more time is available, depending on the degree of erudition of the audience. Chapter VI involves somewhat more complicated techniques, which are absolutely essential to carry on further work in the direction of that chapter. The reader should keep in mind, however, that in order to obtain the most far-reaching and best possible results, it may be necessary to start with a substantially different point of view, and different structure for the proofs, more closely related to what has classically been called diophantine approximations.

In any case, it is remarkable that a mathematical theory as old as the theory of transcendental numbers (dating back to Hermite's first result of 1873, the transcendence of $e$) is still in what can only be called an underdeveloped state.

*Berkeley, 1966* SERGE LANG

# Contents

## Chapter V

### Finitely Generated Values



## Chapter VI

### Transcendence Measures



## Chapter VII

### Linear Differential Equations

CHAPTER I

# Preliminaries

## §1. *Algebraic integers*

A *finite* extension of the field of rational numbers is called a *number field*. According to this convention, the field of all algebraic numbers will not be called a number field.

Let $K$ be a number field. An element $\alpha \in K$ is called an *algebraic integer* if it satisfies either one of the following two equivalent conditions:

INT 1. *There exists a polynomial*

$$f(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0$$

*with* $n \geqq 1$ *and coefficients* $a_i \in \mathbf{Z}$ *such that* $f(\alpha) = 0$.

INT 2. *There exists a finitely generated* $\mathbf{Z}$*-module* $M \neq 0$ (*contained in some algebraic extension of* $K$) *such that* $\alpha M \subset M$.

The equivalence of these conditions is easily proved. Assume INT 1. Let $M$ be generated by $1, \alpha, \ldots, \alpha^{n-1}$. Then $\alpha M \subset M$. Conversely, assume INT 2, and say $M$ is generated by $v_1, \ldots, v_n$. Then

$$\alpha v_i = \sum_{j=1}^{n} a_{ij} v_j$$

for some integers $a_{ij}$. Subtracting the left-hand side from the right, we conclude that the determinant

$$\begin{vmatrix} a_{11} - \alpha & & & a_{ij} \\ & a_{22} - \alpha & & \\ & & \ddots & \\ a_{ij} & & & a_{nn} - \alpha \end{vmatrix}$$

annihilates the module $M$, whence this determinant must be equal to 0. In this way we obtain a polynomial with integer coefficients, leading coefficient 1, which has $\alpha$ as a root.

From condition INT 2, we see that *the set of algebraic integers in $K$ is a ring.* Indeed, if $\alpha$, $\alpha'$ are algebraic integers in $K$, and $M$, $M'$ are finitely generated (non-zero) $\mathbf{Z}$-modules in some algebraic extension of $K$ such that $\alpha M \subset M$ and $\alpha' M' \subset M'$, then $MM'$ is finitely generated, and is mapped into itself by multiplication with $\alpha + \alpha'$ and $\alpha\alpha'$. We denote the ring of algebraic integers in $K$ by $I_K$.

Let $\alpha$ be any element of $K$, satisfying an irreducible equation (over $\mathbf{Z}$)

$$a_m X^m + \cdots + a_0 = 0$$

with $a_i \in \mathbf{Z}$. We assume that the coefficients $a_0, \ldots, a_m$ are relatively prime, and $a_m > 0$. Then this equation is uniquely determined by $\alpha$. A positive integer $d$ will be called a *denominator* for $\alpha$ if $d\alpha$ is an algebraic integer. It is clear that $a_m$ is a denominator for $\alpha$, because if we multiply the above equation by $a_m^{m-1}$ we find that $a_m\alpha$ satisfies the equation

$$(a_m\alpha)^m + a_{m-1}(a_m\alpha)^{m-1} + \cdots + a_m^{m-1}a_0 = 0,$$

with integer coefficients, and leading coefficient 1.

In particular, we see that $K$ is the quotient field of $I_K$, and that every element of $K$ can be written as a quotient of an algebraic integer and a positive (rational) integer.

Each embedding $\sigma: K \to \mathbf{C}$ of $K$ into the complex numbers will be called a *conjugate* of $K$. If $\alpha \in K$, then we call $\sigma\alpha$ a *conjugate* of $\alpha$. The number of conjugates of $K$ is equal to the degree $[K : \mathbf{Q}]$ (dimension of $K$ as vector space over $\mathbf{Q}$). This is a simple elementary fact of field theory.

Let $[K : \mathbf{Q}] = n$. We can map $I_K$ into $\mathbf{C}^n$ by $\tau: \alpha \mapsto (\sigma_1\alpha, \ldots, \sigma_n\alpha)$. This is an additive embedding. In any bounded region of $\mathbf{C}^n$ $(= \mathbf{R}^{2n})$ there is only a finite number of elements of $\tau(I_K)$. Indeed, if we bound a certain region, then we bound the conjugates $\sigma_i\alpha$ of elements $\alpha$ in $I_K$. Each such $\alpha$ is a root of the polynomial

$$(X - \sigma_1\alpha) \cdots (X - \sigma_n\alpha) = X^n + a_{n-1}X^{n-1} + \cdots + a_0,$$

whose coefficients are integers (because they are algebraic integers, and rational numbers being symmetric in the conjugates). Consequently, it follows by an elementary result that *$I_K$ is a free abelian group, whose rank must be precisely $n$* since $K$ is the quotient field of $I_K$, and a basis for $I_K$ over $\mathbf{Z}$ must at the same time be a basis of $K$ over $\mathbf{Q}$. (For a proof of the elementary result, which is standard, cf. for instance my book *Linear Algebra.*)

If $B$ is real $> 0$, we shall say that size$(\alpha) \leqq B$ if there exists a denominator $d$ for $\alpha$ such that $\log d \leqq B$, and if

$$\log \max_{\sigma} |\sigma\alpha| \leqq B$$

for all embeddings $\sigma$ of $K$ into $\mathbf{C}$ (i.e. all conjugates of $\alpha$ have logs of absolute value bounded by $B$). Thus

$$\operatorname{size}(\alpha) = \max(\log d, \log |\sigma\alpha|),$$

where $d$ is the smallest denominator for $\alpha$, and $\sigma$ ranges over all embeddings of $K$ into $\mathbf{C}$. If $r$ is a positive integer, then $\operatorname{size}(\alpha^r) \leqq r \cdot \operatorname{size}(\alpha)$.

Let $\alpha \in K$ and assume $\alpha \neq 0$. Let $d$ be a denominator for $\alpha$. Then $d\alpha$ is an algebraic integer. If $\sigma_1, \ldots, \sigma_n$ are the distinct embeddings of $Q(\alpha)$ into $\mathbf{C}$, then the norm of $d\alpha$ satisfies the inequality

$$1 \leqq |\mathbf{N}(d\alpha)| = \prod_{i=1}^{n} |\sigma_i(d\alpha)| = d^n \prod_{i=1}^{n} |\sigma_i(\alpha)|,$$

because the norm of $d\alpha$ is an algebraic integer and a rational number, whence an ordinary integer, $\neq 0$. From this we obtain the fundamental inequality, to be used many times in this book,

$$\boxed{-[K : \mathbf{Q}] \operatorname{size}(\alpha) \leqq \log |\sigma\alpha|}$$

for any embedding $\sigma$ of $K$ into $\mathbf{C}$. One could obviously refine this inequality, to

$$-[K : \mathbf{Q}] \log \operatorname{den}(\alpha) - ([K : \mathbf{Q}] - 1) \operatorname{size}(\alpha) \leqq \log |\sigma\alpha|.$$

## §2. *Integral linear equations*

We shall prove lemmas due to Siegel, which are used constantly in the sequel. They show that under certain circumstances, one can solve homogeneous linear equations with integer coefficients by means of a solution whose size is approximately the same as the size of the coefficients.

LEMMA 1. *Let*

$$\begin{gathered} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ \cdots \\ a_{r1}x_1 + \cdots + a_{rn}x_n = 0 \end{gathered}$$

*be a system of linear equations with integer coefficients $a_{ij}$, and $n > r$. Let $A$ be a number $\geqq 1$ such that $|a_{ij}| \leqq A$ for all $i, j$. Then there exists an integral, non-trivial solution with*

$$|x_j| \leqq 2(2nA)^{r/(n-r)}.$$

*Proof.* We view our system of linear equations as a linear equation $L(X) = 0$, where $L$ is a linear map, $L\colon \mathbf{Z}^{(n)} \to \mathbf{Z}^{(r)}$, determined by the matrix of coefficients. If $B$ is a positive number, we denote by $\mathbf{Z}^{(n)}(B)$ the

set of vectors $X$ in $\mathbf{Z}^{(n)}$ such that $|X| \leqq B$ (where $|X|$ is the maximum of the absolute values of the coefficients of $X$). Then $L$ maps $\mathbf{Z}^{(n)}(B)$ into $\mathbf{Z}^{(r)}(nBA)$. The number of elements in $\mathbf{Z}^{(n)}(B)$ is $\geqq B^n$ and $\leqq (2B)^n$. We seek a value of $B$ such that there will be two distinct elements $X$, $Y$ in $\mathbf{Z}^{(n)}(B)$ having the same image, $L(X) = L(Y)$. For this, it will suffice that $B^n > (nBA)^r$, and thus it will suffice that $B = (2nA)^{r/(n-r)}$. We take $X - Y$ as the solution of our problem.

The trace Tr from $K$ to $\mathbf{Q}$ establishes an isomorphism between $K$ (as vector space over $\mathbf{Q}$) and its dual space, under the bilinear map $(x, y) \mapsto \mathrm{Tr}(xy)$. Indeed, since the trace is a non-zero linear functional, the kernel on the right and left of this bilinear map is 0.

Let $X = (x_1, \dots, x_n)$ be a vector of elements of $K$. We write

$$\|X\| = \max_{i,\sigma}|\sigma x_i|,$$

that is $\|X\|$ is the maximum of the absolute values of all conjugates of the coordinates $x_i$.

Let $\omega_1, \dots, \omega_M$ be a basis of $I_K$ over $\mathbf{Z}$. Let $\alpha \in I_K$, and write

$$\alpha = a_1\omega_1 + \cdots + a_M\omega_M, \qquad a_j \in \mathbf{Z}.$$

Let $\omega_1', \dots, \omega_M'$ be the dual basis of $\omega_1, \dots, \omega_M$ with respect to the trace. Then we can express the (Fourier) coefficients $a_j$ of $\alpha$ as a trace,

$$a_j = \mathrm{Tr}(\alpha\omega_j').$$

The trace is a sum over the conjugates. Hence the order of magnitude of these coefficients is bounded by that of $\alpha$, times a fixed constant, depending on the elements $\omega_j'$.

Lemma 2. *Let $K$ be a finite extension of $\mathbf{Q}$. Let*

$$\begin{gathered} \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n = 0 \\ \cdots \\ \alpha_{r1}x_1 + \cdots + \alpha_{rn}x_n = 0 \end{gathered}$$

*be a system of linear equations with coefficients in $I_K$, and $n > r$. Let $A$ be a number such that $\|\alpha_{ij}\| \leqq A$, for all $i$, $j$. Then there exists a non-trivial solution $X$ in $I_K$ such that*

$$\|X\| \leqq C_1(C_2nA)^{r/(n-r)} + C_1,$$

*where $C_1$, $C_2$ are constants depending only on $K$.*

*Proof.* Let $\omega_1, \dots, \omega_M$ be a basis of $I_K$ over $\mathbf{Z}$. Each $x_j$ can be written

$$x_j = \xi_{j1}\omega_1 + \cdots + \xi_{jM}\omega_M$$

with unknowns $\xi_{j\lambda}$. Each $\alpha_{ij}$ can be written

$$\alpha_{ij} = a_{ij1}\omega_1 + \cdots + a_{ijM}\omega_M$$

with integers $a_{ij\lambda} \in \mathbf{Z}$. If we multiply out the $\alpha_{ij}x_j$, we find that our linear equations with coefficients in $I_K$ are equivalent to a system of $rM$ linear equations in the $nM$ unknowns $\xi_{j\lambda}$, with coefficients in $\mathbf{Z}$, whose magnitude is bounded by $CA$, where $C$ is a number depending only on $M$ and the size of the elements $\omega_\lambda$, together with the products $\omega_\lambda\omega_\mu$, in other words where $C$ depends only on $K$. Applying Lemma 1, we obtain a solution in terms of the $\xi_{j\lambda}$, and hence a solution $X$ in $I_K$, whose magnitude satisfies the desired bound.

In applying Lemma 2, we meet equations whose coefficients are not necessarily algebraic integers. However, this case is trivially reduced to Lemma 2 by clearing denominators. We formulate it separately.

Lemma 3. *Let $K$ be a number field. Let*

$$\begin{gathered}\alpha_{11}x_1 + \cdots + \alpha_{1n}x_n = 0\\ \cdots\\ \alpha_{r1}x_1 + \cdots + \alpha_{rn}x_n = 0\end{gathered}$$

*be a system of linear equations with coefficients in $K$, and $n > r$. Let $A$ be a number such that $\|\alpha_{ij}\| \leqq A$ for all $i, j$. Let $d_i$ $(i = 1, \ldots, r)$ be a common denominator for the coefficients of the $i$-th equation, and let*

$$d = \max d_i.$$

*Then there exists a non-trivial solution $X$ in $I_K$ such that*

$$\|X\| \leqq C_1(C_2 n\, dA)^{r/(n-r)} + C_1,$$

*where $C_1, C_2$ are constants depending only on $K$.*

*Proof.* We multiply the $i$-th equation by $d_i$, and apply Lemma 2.

It is convenient to mention here the following notation. If

$$P(T) = a_nT^n + \cdots + a_0$$

is a polynomial with complex coefficients $a_i$, we let

$$|P| = \max |a_i|.$$

If the coefficients are algebraic numbers, we let

$$\|P\| = \max \|a_i\|.$$

In *every* application of Lemma 2, we shall deal with a situation when $r$ is approximately equal to $cn$ for some constant $c$ with $0 < c < 1$. Then the expression

$$\frac{r}{n - r}$$

is just equal to $c/(1 - c)$, i.e. is essentially constant, independent of the magnitude of $n$ or $r$ which will be quite large. Thus in this circumstance, the solution is bounded by

$$CnA^{c'}$$

for some constant $c'$. It will also be the case that $n$ is very small compared to $A$ (essentially, $n$ will be of the order of magnitude of $\log A$), and consequently we shall have solutions whose order of magnitude is $A^{c'}$.

## §3. *Estimating*

In all our work, we shall estimate. To do this efficiently, we need various notation.

Let $f$, $g$ be real valued functions defined for arbitrarily large real numbers, and assume $g \geqq 0$. We write

$$f = O(g)$$

if there is a constant $C > 0$ such that $|f(x)| \leqq Cg(x)$ for all sufficiently large $x$. We write

$$f = o(g)$$

if $\lim_{x\to\infty} f(x)/g(x) = 0$.

We write $f \ll g$ if there exists a constant $C > 0$ such that for all $x$ sufficiently large we have $f(x) \leqq Cg(x)$. It will always be made clear from the context what the constant $C$ depends on. If $f$, $g$ are both $\geqq 0$, we write $f \gg\ll g$ to mean $f \ll g$ and $g \ll f$.

In applications, $f$ and $g$ may be defined for all sufficiently large positive integers.

We shall also use the $\ll$ notation when $f$, $g$ are defined on some set $S$ (not a set of numbers). In that case, $f \ll g$ means that there exists a number $C > 0$ such that $f(x) \leqq Cg(x)$ for all elements of the set.

We shall frequently estimate sums. This is always done in a relatively coarse manner, namely the absolute value of a sum is bounded by the number of terms in the sum, times the maximum value of the terms. It will also turn out in each case that the number of terms in the sum is very small compared to the value of the terms (something like the log), so that the number of terms in a sum can always be essentially disregarded in making our estimates.

In estimating sums, one deals most easily with the ordinary triangle inequality. However, all the estimates will be important only as they appear in exponents. Thus it is convenient to take the log to simplify the notation. The reader will no doubt need, as I do occasionally, to reformulate the estimates explicitly in the exponential form, but after having done this several times, he will appreciate the other notation as a means of achieving expository clarity.

Finally, we describe very briefly the kind of estimates which lead to transcendence proofs. They are all based on the following principle, at the present stage of the theory. One deals with entire, or meromorphic functions, say $f$, $g$, which are assumed algebraically independent, over the constant field, and are of finite order, in the analytic sense. It is known from complex variables, that such functions cannot have too many zeros.

We are then interested in those complex numbers $z$ such that $f(z)$ and $g(z)$ are both algebraic. Our functions will usually have some additional property like satisfying an algebraic differential equation, or possessing an algebraic addition theorem, which, out of one such value, allows the generation of many such values, and the problem is to describe the restrictions on those numbers $z$ in such a way as to end up with sharp criteria. The arithmetic problem involving algebraic values is reduced to an analytic one concerning *zeros* of an auxiliary function, a polynomial in $f$, $g$,

$$F = \sum a_{ij} f^i g^j$$

with coefficients in a number field. From the assumption that $f$, $g$ take on values in the number field at certain points, one can construct a function $F$ having many zeros, using Siegel's lemma. By some form of a three circle theorem, one can then obtain a contradiction, since on the one hand, $F$ has small absolute value at some point $w$ where $f(w)$ and $g(w)$ are algebraic, and on the other hand, the construction of $F$ could be achieved in such a way that the size of $F(w)$ was relatively small. The contradiction then comes from the fundamental inequality mentioned in §1, relating the size of an algebraic number with one of its absolute values. Each chapter will exhibit a variation of the general principle just described.

It should be pointed out that Gelfond was the first to realize explicitly the connection between transcendence problems and algebraic values of entire functions. Many years before he proved his $\alpha^\beta$ theorem, investigating special cases in 1929 (cf. [12]) he saw that such a problem was related with a question (I believe raised by Polya) concerning the possibility of an entire function taking integral values at integers, and the interpolation problem arising from it. A function like $F$ above was first introduced by Siegel in connection with the transcendence of values of the Bessel function.

CHAPTER II

# Meromorphic Functions

## §1. *Algebraic values of* $e^t$

The function $e^t$ is the simplest function to consider, and as promised in the foreword, we begin by treating it as a special case of more general functions to be handled later.

We note the trivial fact that

$$\max_{|t|=R} |e^t| \leqq e^R.$$

This kind of growth condition is used all the time.

Let $\beta_1, \ldots, \beta_m$ be complex numbers, linearly independent over the rationals. Then the functions

$$e^{\beta_1 t}, \ldots, e^{\beta_m t}$$

are algebraically independent over the complex numbers. In fact, if we denote these functions by $\chi_1, \ldots, \chi_m$ then we note that they are multiplicatively independent: In any relation

$$\chi_1^{n_1} \cdots \chi_m^{n_m} = 1$$

with integer exponents $n_1, \ldots, n_m$ we must have $n_1 = \cdots = n_m = 0$. From this it follows easily that $\chi_1, \ldots, \chi_m$ are algebraically independent. One may view this as a special case of Artin's theorem on the independence of characters, or the reader may devise a simple proof using his own ingenuity.

THEOREM 1. *Let $\beta_1$, $\beta_2$ be complex numbers, linearly independent over* $\mathbf{Q}$, *and let $z_\nu$ ($\nu = 1, 2, 3$) be complex numbers, also linearly independent over* $\mathbf{Q}$. *Then at least one of the numbers*

$$e^{\beta_1 z_\nu}, e^{\beta_2 z_\nu} \qquad (\nu = 1, 2, 3)$$

*is transcendental (over* $\mathbf{Q}$*).*

Before proving Theorem 1, we give some corollaries.

COROLLARY 1. *Let $\beta$ be a complex number, and suppose that there exist three non-zero, multiplicatively independent algebraic numbers*

$$\alpha_\nu \quad (\nu = 1, 2, 3)$$

*such that $\alpha_\nu^\beta$ is algebraic. Then $\beta$ is rational.*

*Proof.* Suppose that $\beta$ is irrational. Let $\beta_1 = 1$, $\beta_2 = \beta$, and $z_\nu = \log \alpha_\nu$ (with any determination of the logarithm). The corollary follows by a direct application of Theorem 1.

COROLLARY 2. *Let $y$ be real, and $x^y$ algebraic for all positive rational $x \neq 0$. Then $y$ is rational.*

*Proof.* Special case of Corollary 1.

We view the function $e^t$ as an entire function, i.e. a function which is analytic in the plane.

We recall that an entire function $F$ is said to be of *order* $\leqq \rho$ if there is a constant $C > 0$ such that

$$|F|_R = \max_{|t|=R} |F(t)| \leqq C^{R^\rho}$$

for all $R$ sufficiently large. (We omit the usual $\epsilon$ which the reader will find in texts on complex variables, because it is irrelevant for what follows.) With the notation of Chapter I, §3, we may write

$$\log |F|_R = O(R^\rho)$$

or also

$$\log |F|_R \ll R^\rho$$

for $R$ sufficiently large. The implied constant then depends on $F$. In all books on complex variables, it is proved that the number of zeros of such a function $F$ in a circle of radius $R$ is $O(R^\rho)$, provided $F$ is not identically zero.

We shall now prove Theorem 1. Suppose that the conclusion of Theorem 1 is false, and let $K$ be a finite extension of $\mathbf{Q}$, containing

$$e^{\beta_1 z_\nu},\ e^{\beta_2 z_\nu}$$

for $\nu = 1, 2, 3$. Let $n$ be a large integer, assumed to be a square for convenience, which will tend to infinity later. Let $r = (4n)^{3/2}$. We can find algebraic integers $a_{ij}$ not all zero in $K$ such that the function

$$F(t) = \sum_{i,j=1}^{r} a_{ij} e^{i\beta_1 t} e^{j\beta_2 t}$$

has a zero at every point $k \cdot z = k_1z_1 + k_2z_2 + k_3z_3$, where $k = (k_1, k_2, k_3)$ is a triple of integers such that $1 \leqq k_\nu \leqq n$. This amounts to solving linear equations in $r^2$ unknowns, with $r^2 = (4n)^3$, and

$$\text{number of equations} = n^3.$$

The coefficients of the equations are the values

$$e^{i\beta_1(k\cdot z)}e^{j\beta_2(k\cdot z)}$$

which are elements of $K$. For each equation (corresponding to some $k$), we have

$$\text{size of coefficients} \ll nr \ll n^{5/2},$$

the implied constant depending on the values $\exp(\beta_\mu z_\nu)$. For each $k$, we can find a common denominator $d$ for the coefficients of the $k$-th equation, satisfying the bound

$$\log d \ll nr,$$

because the coefficients are powers of the fixed algebraic numbers $\exp(\beta_\mu z_\nu)$, and these powers are essentially bounded by $nr$. We can therefore apply Siegel's lemma, and we can in fact find the $a_{ij}$ such that

$$\text{size } a_{ij} \ll n^{5/2}$$

for $n$ sufficiently large.

Since $\beta_1$, $\beta_2$ are linearly independent over $\mathbf{Q}$, it follows that $F$ is not identically zero, and takes on values in $K$ for all linear combinations of $z_1, z_2, z_3$ with positive integer coefficients. On the other hand, $F$ cannot vanish at all such linear combinations, because they are not discrete, or alternatively because $F$ is entire of order 1, and in a circle of large radius $R$, there are more such linear combinations than the bound $O(R)$ for the number of possible zeros of $F$. Let $s$ be the largest integer such that $F(k \cdot z) = 0$ for all $k_\nu$ with $1 \leqq k_\nu \leqq s$. Then $s \geqq n$. Let

$$w = k_1z_1 + k_2z_2 + k_3z_3$$

with some $k_\nu = s + 1$, and $1 \leqq k_\nu \leqq s + 1$ for all $\nu$, and $F(w) \neq 0$. Then

$$\text{size } F(w) \ll s^{5/2}.$$

We now estimate $|F(w)|$, and we use the expression

$$F(w) = \frac{F(t)}{\prod(t - k \cdot z)} \prod (w - k \cdot z)\Big|_{t=w},$$

the products being taken over all $k_\nu$ with $1 \leqq k_\nu \leqq s$. There are $s^3$ terms in the product. The function on the right of this last equality is an entire function, and we apply the maximum modulus principle on a circle of

radius $R = s^{3/2}$. Note that for $|t| = R$, we have $|t - k \cdot z| \geqq R/2$ (for $s$ large), and also

$$\frac{|w - k \cdot z|}{|t - k \cdot z|} \leqq \frac{C_1 s}{R} \leqq \frac{C_1}{s^{1/2}}$$

for some constant $C_1$, and $s$ large. Hence

$$\log |F(w)| \ll \log |F|_R + s^3 - \tfrac{1}{2}s^3 \log s.$$

A trivial estimate shows that

$$|F|_R \leqq r^2 C_2^{n^{5/2}} C_3^{rR} \leqq C_4^{s^3},$$

and hence

$$\log |F(w)| \ll s^3 - s^3 \log s.$$

This contradicts the lower bound

$$-\text{size}\, F(w) \ll \log |F(w)|,$$

if we let $n$, and therefore $s$, tend to infinity. Theorem 1 is proved.

*Remark.* One would like to shrink the number of $z_\nu$ from three to two, but the same pattern of proof in this case does not give the desired contradiction at the end (it just misses).

In investigating values of $\alpha^\beta$ when $\beta$ is transcendental and $\alpha$ is algebraic, there may be one algebraic $\alpha$ such that $\alpha^\beta$ is algebraic. For instance,

$$2^{\frac{\log 3}{\log 2}} = 3.$$

It will be shown later (Gelfond-Schneider Theorem) that $\alpha^\beta$ is transcendental if $\alpha$ is algebraic $\neq 0, 1$ and $\beta$ is irrational. Thus $(\log 3)/(\log 2)$ is transcendental, and constitutes an exceptional number $\beta$ for which $2^\beta$ is algebraic. Theorem 1 shows that there are at most two multiplicatively independent possibilities, and our remark is that there should only be one. In any case, we see Theorem 1 as a complement of the Gelfond-Schneider theorem. It also shows that among the numbers $2^\pi, 3^\pi, 5^\pi, \ldots$ at most two are algebraic. Of course, in this case, one expects none of them to be algebraic.

## §2. *Algebraic values of meromorphic functions*

We shall axiomatize the proof of Theorem 1 so that it applies to more general functions.

A meromorphic function $f$ is said to be of *order* $\leqq \rho$ if it can be expressed as a quotient of entire functions of order $\leqq \rho$. If $S$ is a set of

complex numbers, we say that such a meromorphic function $f$ is *defined* on $S$ if we can write $f = g/h$ where $g$, $h$ are entire of order $\leqq \rho$ and no point of $S$ is a zero of $h$. For simplicity, we shall *always* assume that $\rho$ is an integer $\geqq 1$.

Let $S$ be a set of complex numbers, expressed as a union, $S = \bigcup_{n=1}^{\infty} S_n$, such that $S_n \subset S_{n+1}$ for all $n$. We say that the subsets $\{S_n\}$ form a filtration of $S$. We shall assume throughout when dealing with such filtrations that there is a constant $C > 0$ such that for all $n$ and all $z \in S_n$ we have $|z| \leqq Cn$.

Let $f$ be a meromorphic function defined on $S$, with values in a number field $K$. We shall say that $f$ is of *arithmetic order* $\leqq \rho$ on $S$ (or more accurately, with respect to the filtration $\{S_n\}$) if there is a constant $C \geqq 1$ such that the following conditions are satisfied:

AO 1. *For all $n$ and $z \in S_n$ we have size $f(z) \leqq Cn^\rho$.*

AO 2. *There is an entire function $h$ of order $\leqq \rho$, such that $hf$ is entire, $h$ has no zero in $S$, and for all $n$, $z \in S_n$,*

$$\log |1/h(z)| \leqq Cn^\rho.$$

Using our notation $\ll$, we can write our conditions in the form

$$\operatorname{size} f(z) \ll n^\rho \qquad \text{and} \qquad \log |1/h(z)| \ll n^\rho,$$

for $z \in S_n$, and $n \to \infty$.

THEOREM 2. *Let $f$, $g$ be meromorphic functions of order $\leqq \rho$. Let $S = \bigcup S_n$ be as above, and assume that $f$, $g$ are defined on $S$, with values in the number field $K$, and of arithmetic order $\leqq \rho$ on $S$. Let $\lambda$ be a number $> 2$. If* $\operatorname{card}(S_n) \gg\ll n^{\lambda\rho}$ *for $n \to \infty$, then $f$, $g$ are algebraically dependent over $K$.*

*Proof.* Let $C_1, C_2 > 0$ be such that

$$C_1 n^{\lambda\rho} \leqq \operatorname{card}(S_n) \leqq C_2 n^{\lambda\rho}$$

for all $n$. Without loss of generality, we may assume (for convenience) that $C_2$ is an integer, and is a square. We shall deal with a large integer $n$, taken to be a square, and which will tend to infinity later.

Let $r = 2C_2^{1/2} n^{\lambda\rho/2} \ll n^{\lambda\rho/2}$.

We can find algebraic integers $a_{ij}$ not all zero in $K$, such that the function

$$F = \sum_{i,j=1}^{r} a_{ij} f^i g^j$$

has a zero at every point $z \in S_n$. This amounts to solving linear equations in $r^2$ unknowns, and we have

$$\text{number of equations} \leqq C_2 n^{\lambda\rho} \leqq r^2/4.$$

The coefficients of these equations are the values

$$f(z)^i g(z)^j$$

with $z \in S_n$. By hypothesis, we have

$$\text{size of coefficients} \ll n^\rho r.$$

For each $z$, we also note that there is a common denominator for the coefficients of the corresponding equation with the similar bound $\ll n^\rho r$. By Siegel's lemma, we can find the $a_{ij}$ such that

$$\text{size } a_{ij} \ll n^\rho r$$

for all $n$ sufficiently large.

If $f$, $g$ are algebraically independent over $K$, then $F$ is not identically zero. Let $s$ be the smallest integer such that $f(z) = 0$ for all $z \in S_s$ but $F(w) \neq 0$ for some $w \in S_{s+1}$. (Such $s$ exists because in a circle of radius $s$ we know that $F$ has $O(s^\rho)$ zeros for $s \to \infty$.) Then $s \geqq n$ and trivial estimates yield

$$\text{size } F(w) \ll s^\rho r.$$

We note that $s^\rho r \ll s^{\lambda\rho}$.

We now estimate $|F(w)|$. Let $h$ be the entire function satisfying condition AO 2 with respect to both $f$ and $g$. (If $h_1$ satisfies this condition with respect to $f$, and $h_2$ satisfies it with respect to $g$, then $h = h_1 h_2$ will serve our purposes.) Then $h^{2r}F$ is an entire function having zeros at all elements of $S_s$. We use the expression

$$F(w) = \frac{h(t)^{2r}F(t)}{h(w)^{2r}\prod(t - z)} \prod (w - z)\bigg|_{t=w}$$

where the products are taken for $z$ in $S_s$. We apply the maximum modulus principle, and bound the function on the right on a circle of radius $R = s^{\lambda/2}$. We estimate three things:

First, we estimate $h^{2r}F$ using the fact that $h$, $hf$, $hg$ are entire of order $\leqq \rho$, and using the expression

$$h^{2r}F = \sum_{i,j=1}^{r} a_{ij} h^{2r-i-j}(hf)^i(hg)^j.$$

We take the log of the absolute value, and find

$$\sup_{|t|=R} \log |h(t)^{2r}F(t)| \ll s^{\lambda\rho/2}r \ll s^{\lambda\rho}.$$

Second, we have the upper bound (by AO 2),

$$\log |1/h(w)^{2r}| \ll s^{\rho}r \ll s^{\lambda\rho}.$$

Third, and most important, we have

$$\log \sup_{|t|=R} \prod_{z \in S_s} \frac{|w - z|}{|t - z|} \ll s^{\lambda\rho} - \frac{\lambda - 2}{2} s^{\lambda\rho} \log s.$$

Combining these three estimates, we get the upper bound

$$\log |F(w)| \ll s^{\lambda\rho} - \frac{\lambda - 2}{2} s^{\lambda\rho} \log s$$

(because $3s^{\lambda\rho} \ll s^{\lambda\rho}$!). For $s$ sufficiently large, this contradicts the fact that

$$-\text{size}\, F(w) \ll \log |F(w)|,$$

thereby proving Theorem 2.

*Remark.* Theorem 2 remains valid if instead of assuming

$$\text{card}(S_n) \gg\ll n^{\lambda\rho}$$

we assume merely that $\text{card}(S_n) \gg n^{\lambda\rho}$. Indeed, in that case, we reduce the proof to the preceding one by constructing sets $S'_n \subset S_n$ such that $\text{card}(S'_n) \gg\ll n^{\lambda\rho}$ and $S'_n \subset S'_{n+1}$. This is trivially done inductively, and we can then apply Theorem 2 to the sets $S'_n$.

## §3. *Application to linear groups*

We shall now assume that the reader is acquainted with group varieties, but will recall briefly the main properties of this type of object.

A group variety is a group in affine or projective space, which is also an algebraic variety (connected), i.e. its points are the set of solutions of algebraic equations, and such that the law of composition and inverse have graphs which are also algebraic varieties. An important example is the linear group $GL(m)$ of invertible $m \times m$ matrices.

If $K$ is a subfield of the complex numbers, we say that the group variety is defined over $K$ if all the above-mentioned algebraic equations can be chosen to have coefficients in $K$. If that is the case, then we denote by $G_K$ the set of points of $G$ having coordinates in $K$, and it follows that $G_K$

is a group. When $K$ is a number field, we view $G_K$ as a discrete group. When $K = \mathbf{C}$, we view $G_{\mathbf{C}}$ as a complex-analytic manifold, i.e. a complex analytic group.

Let $G$ be a group variety. By a *1-parameter subgroup* of $G$ we mean a complex-analytic homomorphism $\varphi: \mathbf{C} \to G_{\mathbf{C}}$ of the complex line into $G_{\mathbf{C}}$ whose derivative at the origin is injective. Thus $\varphi$ is an analytic curve in $G_{\mathbf{C}}$. However, it may well be that algebraically, $\varphi$ has dimension $> 1$. Indeed, if, say, $G$ is embedded in some projective space, and $(\varphi_0, \dots, \varphi_N)$ are the projective coordinates of $\varphi$, then by the *algebraic dimension* of $\varphi$ we shall mean the maximum number of algebraically independent coordinate functions $\varphi_i/\varphi_0$ $(i = 1, \dots, N)$. This is the same as the dimension of the smallest group subvariety of $G$ containing $\varphi(\mathbf{C})$, i.e. containing the 1-parameter subgroup.

THEOREM 3. *Let $G$ be a linear group variety defined over the field of algebraic numbers. Let $\varphi: \mathbf{C} \to G_{\mathbf{C}}$ be a 1-parameter subgroup, of algebraic dimension $\geqq 2$. Let $\Gamma$ be a subgroup of $\mathbf{C}$ having at least three linearly independent elements over $\mathbf{Q}$. Then $\varphi(\Gamma)$ cannot be contained in the group of algebraic points of $G$.*

*Proof.* We can represent $G$ globally as a group of matrices, so that for some matrix $M$,

$$\varphi(t) = \exp(tM) = \sum t^\mu M^\mu/\mu!.$$

Let $z_\nu$ $(\nu = 1, 2, 3)$ be complex numbers linearly independent over $\mathbf{Q}$, such that $\varphi(z_\nu)$ is algebraic, say with components in a number field $K$. Suppose that $M$ is an $m \times m$ matrix. Then we can write

$$\varphi(t) = (\varphi_{ij}(t)), \qquad 1 \leqq i, j \leqq m,$$

with entire functions $\varphi_{ij}$ which are of order $\leqq 1$. If $\varphi$ has algebraic dimension $\geqq 2$, then at least two among the functions $\varphi_{ij}$ are algebraically independent over $K$, say $f$, $g$.

Let $S_n$ be the set of linear combinations

$$k_1 z_1 + k_2 z_2 + k_3 z_3$$

with $1 \leqq k_\nu \leqq n$. Condition AO 1 is easily verified, because

$$\varphi(k_1 z_1 + k_2 z_2 + k_3 z_3) = \varphi(z_1)^{k_1} \varphi(z_2)^{k_2} \varphi(z_3)^{k_3},$$

and it is easy to make the necessary estimates on the size of a product of matrices to see that AO 1 is true. Condition AO 2 is always true for entire functions (take $h = 1$), and consequently, we can apply Theorem 2 to obtain the desired contradiction.

COROLLARY. *Let $G$ be a linear group variety defined over a number field $K$. Let $\Gamma$ be a commutative subgroup of $G_K$, containing at least three elements independent over $\mathbf{Z}$. If $\Gamma$ is contained in a 1-parameter subgroup of $G$, then this 1-parameter subgroup is a group subvariety (i.e. algebraic) of dimension 1.*

We observe that Theorem 1 of §1 is a special case of Theorem 3. Indeed, it involves the 1-parameter subgroup

$$t \mapsto (e^{\beta_1 t}, e^{\beta_2 t})$$

of the product of the multiplicative group with itself, $\mathbf{C}^* \times \mathbf{C}^*$.

We note that the additive group of $m \times m$ matrices may be viewed as the tangent space at the origin of the general linear group $GL(m)$. Thus the map

$$\exp\colon \mathrm{Mat}_m(\mathbf{C}) \to GL(m, \mathbf{C})$$

given by the exponential power series is the usual exponential map of the differential geometer. When the group variety is the multiplicative group $\mathbf{C}^*$, then the exponential map is simply the ordinary function $t \mapsto e^t$. A 1-parameter subgroup is nothing but the image of a straight line through the origin under the exponential map.

In the next section, we consider another type of group variety.

## §4. *Abelian varieties*

An abelian variety is a group variety in projective space. If $A$ is an abelian variety, then $A_{\mathbf{C}}$ is a compact, complex Lie group, and consequently is a complex torus: We can represent $A_{\mathbf{C}}$ as a factor group of $\mathbf{C}^d$ ($d = \dim A$) by a discrete subgroup of dimension $2d$ over $\mathbf{R}$. This representation can be achieved by theta functions in $d$ variables, these being entire functions of order $\leqq 2$ (cf. [3] or [32]). We shall denote this representation by

$$\Theta\colon \mathbf{C}^d \to A_{\mathbf{C}}.$$

If $\alpha \in \mathbf{C}^n$, and $\alpha \neq 0$, then the curve $\varphi\colon \mathbf{C} \to A_{\mathbf{C}}$ such that $\varphi(t) = \Theta(t\alpha)$ is a 1-parameter subgroup of $A$, and all 1-parameter subgroups of $A$ can be so described. Furthermore, $\Theta$ is a representation of the exponential map on $A_{\mathbf{C}}$, and we may identify the tangent space at the origin of $A_{\mathbf{C}}$ with $\mathbf{C}^d$.

THEOREM 4. *Let $A$ be an abelian variety defined over the field of algebraic numbers. Let $\varphi\colon \mathbf{C} \to A_{\mathbf{C}}$ be a 1-parameter subgroup of $A$. Let $\Gamma$ be a subgroup of $\mathbf{C}$ having at least seven linearly independent elements over $\mathbf{Q}$. If $\varphi$ has algebraic dimension $\geqq 2$, then $\varphi(\Gamma)$ cannot be contained in the group of algebraic points of $A$.*

*Proof.* Let $\Theta = (\theta_0, \ldots, \theta_N)$ be the representation of $\Theta$ by means of theta functions $\theta_0, \ldots, \theta_N$, giving an embedding of $A_{\mathbf{C}}$ into projective $N$-space. We induce the functions on the analytic curve $\varphi$, and obtain entire functions $(f_0, \ldots, f_N)$ in one variable, of order $\leqq 2$, which realize our 1-parameter map $\varphi$, that is $\varphi(t) = (f_0(t), \ldots, f_N(t))$. Without loss of generality, we may assume that the divisor of zeros of $\theta_0$ does not pass through the origin, and that, in particular, $f_0$ is not identically zero. Let $D$ be a small disc of radius $\delta > 0$ around the origin in $\mathbf{C}$, so that no zero of $f_0$ lies in $D$. Then $|f_0(t)|$ is bounded away from 0 on $D$, and hence $\log |1/f_0(t)|$ is bounded from above on $D$, thereby satisfying condition AO 2.

By hypothesis, at least two of the functions among $f_1/f_0, \ldots, f_N/f_0$ are algebraically independent over the algebraic numbers, say $f, g$.

We may assume that $A$ is defined over a number field $K$, and that $\Gamma$ is generated by seven elements $z_\nu$ $(\nu = 1, \ldots, 7)$ linearly independent over $\mathbf{Q}$, such that $\varphi(z_\nu)$ is contained in $A_K$ for $\nu = 1, \ldots, 7$. Let $S_n$ be the set of linear combinations

$$k \cdot z = k_1 z_1 + \cdots + k_7 z_7$$

with $-n \leqq k_\nu \leqq n$, and such that $k \cdot z$ lies in $D$. By routine techniques (which will be recalled below), one verifies that $\mathrm{card}(S_n) \gg n^5$ or $n^6$, depending on whether at least two of the elements $z$ are linearly independent over $\mathbf{R}$ or not. For definiteness, let us assume that we are in the case $\mathrm{card}(S_n) \gg n^5$. Let $S = \bigcup S_n$. We already know that our two functions $f, g$ satisfy condition AO 2 on $S$, since in fact they have the bounded denominator $h = f_0$ on $D$. We shall prove in the next lemma that they satisfy condition AO 1, and thus are of arithmetic order $\leqq 2$ on $S$. Applying Theorem 2 then gives the contradiction which proves Theorem 4.

LEMMA 1. *Let $A$ be an abelian variety defined over a number field $K$. Let $\varphi: \mathbf{C} \to A_{\mathbf{C}}$ be a 1-parameter subgroup, represented by projective coordinates $(f_0, \ldots, f_N)$. Say $f_0$ is not identically zero. Let $z_1, \ldots, z_m \in \mathbf{C}$ be linearly independent over $\mathbf{Q}$, such that $\varphi(z_\nu) \in A_K$ for $\nu = 1, \ldots, m$. Let $S_n^*$ be the set of linear combinations*

$$k_1 z_1 + \cdots + k_m z_m, \qquad -n \leqq k_\nu \leqq n,$$

*and let $Z_n$ be the subset of $S_n^*$ consisting of the zeros of $f_0$. Let $S_n$ be the complement of $Z_n$ in $S_n^*$. Let $S = \bigcup S_n$. Let $f = f_i/f_0$ for some index $i = 1, \ldots, N$. Then $f$ satisfies condition* AO 1 *on $S$, with $\rho = 2$.*

*Proof.* We shall use the quadratic form of Néron-Tate [25], [20]. We denote by $h$ the logarithmic height, defined on the group of rational points $A_K$ of $A$ in $K$. Then $h = q + l + O(1)$, for some quadratic form $q$, and

some linear form $l$. If we write $P_\nu = \varphi(z_\nu)$, and

$$P = k_1P_1 + \cdots + k_mP_m = \varphi(k_1z_1 + \cdots + k_mz_m)$$

with $-n \leqq k_\nu \leqq n$, then we see that

$$(1) \qquad h(P) \ll n^2 \quad \text{for} \quad n \to \infty.$$

What we need is to bound the size of a single coordinate function $f(z)$ for $z \in S_n$ in terms of the height. This is a trivial technical matter. Indeed, the height of a point $P = (\xi_0, \ldots, \xi_N)$ in projective space over $K$ is defined by

$$h(P) = \sum_v \sup_j v(\xi_j)$$

where the sum is taken over the set of normalized absolute values on $K$, and $v(\xi_j) = \log \|\xi_j\|_v$ (notation as in [18], Chapter III, except that we take the logarithm). From this definition, it is clear that if, say $\xi_0 \neq 0$, and $\alpha = \xi_i/\xi_0$ is the $i$-th affine coordinate of our point, then

$$(2) \qquad h(\alpha) = h(\xi_0, \xi_i) \leqq h(\xi).$$

Let $d(\alpha)$ be the leading coefficient of the irreducible polynomial satisfied by $\alpha$ over the *integers* $\mathbf{Z}$, with relatively prime coefficients. Then it is well known that $\log d(\alpha) \ll h(\alpha)$, for $\alpha \in K$. (See for instance [18], Proposition 4 of Chapter III, §2. The implied constant depends on the degree of $\alpha$ over $\mathbf{Q}$.) Furthermore, $d(\alpha)$ is a denominator for $\alpha$, that is $d(\alpha)\alpha$ is integral over $\mathbf{Z}$. From this and the estimate for the height, we see that

$$(3) \qquad \text{size}(\alpha) \ll h(\alpha)$$

for $\alpha \in K$. Putting (1), (2), (3) together, we find that $f$ satisfies condition AO 1 on $S$ with $\rho = 2$, as was to be shown.

There remains but to justify that the number of linear combinations $k \cdot z$ lying in $D$ has at least the order of magnitude of $n^5$. This is done by standard arguments. Suppose for instance that $z_6, z_7$ are linearly independent over $\mathbf{R}$, and let $\Lambda$ be the lattice generated by them. We consider the combinations

$$k_1z_1 + \cdots + k_5z_5$$

on the torus $\mathbf{C}/\Lambda = \mathbf{R}^2/\Lambda$. We cut up a fundamental domain into approximately $1/\delta$ small squares of sides approximately equal to $\delta^{1/2}$, and use the Dirichlet box principle on the $(cn)^5$ elements consisting of linear combinations $k \cdot z$ with $|k| \leqq cn$ for some constant $c$ depending on the $z_\nu$. Subtracting one element from all the others crowded in a small box, we obtain essentially $\gg \delta n^5$ elements in $D$ (up to a constant factor, depending only on the $z_\nu$). This does what we wanted.

(By being slightly more careful, and using translations by $z_6$, $z_7$, using points lying outside small circles around the zeros of the entire function $f_0$, and using the minimum modulus principle for entire functions, one can probably avoid the drop of 7 to 5, and thus replace 7 by 5 in the theorem. However, this is a secondary matter here.)

COROLLARY 1. *Let $A$ be an abelian variety defined over a number field $K$. Let $\Gamma$ be a finitely generated subgroup of $A_K$, of rank $\geqq 7$. If $\Gamma$ is contained in a 1-parameter subgroup of $A_{\mathbf{C}}$, then this 1-parameter subgroup is an abelian subvariety of dimension 1 (an elliptic curve) in $A$.*

COROLLARY 2. *Let $G$, $G'$ be group varieties defined over a number field. Let $G$ have dimension 1, and let $G'$ be linear or abelian. Let $\varphi\colon G_{\mathbf{C}} \to G'_{\mathbf{C}}$ be a complex analytic homomorphism. If there exist 7 algebraic points $\alpha_\nu$ ($\nu = 1, \ldots, 7$) on $G$, linearly independent over $\mathbf{Z}$, such that $\varphi(\alpha_\nu)$ is algebraic, then some power of $\varphi$ is a rational homomorphism.*

*Proof.* Composing $\varphi$ with the exponential map on $G$ shows that the graph of $\varphi$ is algebraic. That some power of $\varphi$ is rational then follows from trivial facts about group varieties.

*Remark.* In the study of abelian varieties, certain theorems, like the Mordell-Weil theorem, and Siegel's theorem concerning integral points on curves, were originally proved using theta functions. However, these theorems dealt only with the algebraic-arithmetic aspects of the situation, and when an algebraic theory of abelian varieties was developed (by Weil), it became clear that one could give expositions for the proofs entirely within the algebraic context. (Néron did it for the Mordell-Weil theorem in his thesis, and I did it for the Siegel theorem, cf. [18].) In other words, the theta functions were used only as a convenient tool to write down the group law on abelian varieties, and this tool became superfluous for the preceding applications when one saw how to formulate the group law purely algebraically.

Besides purely algebraic theorems on abelian varieties, there are other types, for instance those concerned with the complex analytic structure (as in Weil's book on Kähler manifolds). In the present context, we deal still with a third aspect of abelian varieties, namely the direct study of the properties of their transcendental parametrizations, with an arithmetic point of view, in which this parametrization occupies a central position.

## *Historical note*

The use of a function similar to our $F$ occurs in Gelfond's and Schneider's proof of the transcendence of $\alpha^\beta$ when $\alpha$ is algebraic $\neq 0, 1$ and $\beta$ is algebraic irrational. Theorem 1 was known to Siegel [cf. L. Alaoglu and P. Erdös, "On highly composite and similar numbers", *Trans. Am. Math.*

*Soc.* **56** (1944) p. 455]. Corollary 2 is the portion of Theorem 1 needed for applications. Its need had arisen in the paper just mentioned, and more recently was asked by Serre in connection with the determination of certain characters of idele classes of number fields. Unfortunately, Theorem 1 does not appear in any of the three books on transcendental numbers [13], [29], [32] and is not otherwise generally known. Theorem 2 is a variant of results of Schneider [28], who started to axiomatize the transcendence proofs, and showed more explicitly than Gelfond how the order of growth of an entire function restricts its algebraic values.

Further comments on the exponential function will be made at the end of the next chapter. The applications of §3 and §4 are taken from [17].

It is a problem to extend Theorem 2 and its applications to functions of several variables. Formally, the proof can be given a $d$-variable version, the only difficulty occurring at the estimate of an entire function with many zeros. One needs an analogue to the three circle theorem, i.e. an estimate of type

$$|F|_R \leqq \frac{|F|_{R_2} R^N}{R_2^N}$$

if $F$ has $N$ zeros more or less evenly distributed inside a polydisc of radius $R_1$, and $R_1 < R < R_2$. That the zeros must be more or less evenly distributed is clear, since a function of several variables always has a divisor of zeros (i.e. the zeros are not isolated). Consequently, in the multidimensional analogue of Theorems 3 and 4, taking linear combinations of vectors, it is necessary to assume that these combinations are evenly distributed. In applications, this leads to questions in diophantine approximations which at the moment appear to lie beyond the methods available. (For a special case when this problem does not arise, cf. Chapter IV.)

Finally, it would be worth while to extend Theorems 3 and 4 to arbitrary group varieties (which one may assume commutative, since the Zariski closure of the 1-parameter group is necessarily commutative). This is not a problem in transcendental numbers, since Theorem 2 reduces the question to general properties on group varieties. Thus one must show that Axioms AO 1 and AO 2 are satisfied. This involves first a purely algebraic question, i.e. an upper bound for the height of points on arbitrary commutative group varieties (to replace the upper bound derived from the presently available Néron-Tate form), and second, a mixed algebraic-transcendental problem, i.e. showing that the exponential map on arbitrary commutative group varieties can be parametrized by meromorphic functions of order $\leqq 2$, similar to theta functions. (This would involve systematizing Severi's quasi-abelian functions.)

CHAPTER III

# Algebraic Differential Equations

## §1. The main theorem

In the preceding chapter, we saw that there exist certain restrictions on the set of numbers at which algebraically independent meromorphic functions take on algebraic values. In the applications, we use very strongly the algebraic addition formula for our functions. In this chapter, we shall describe an analogous situation when the functions satisfy an algebraic differential equation. In that case, the set of points is much smaller.

Theorem 1. *Let $K$ be a number field. Let $f_1, \dots, f_N$ be meromorphic functions of order $\leqq \rho$. Assume that the field $K(f_1, \dots, f_N)$ has transcendence degree $\geqq 2$ over $K$, and that the derivative $D = d/dt$ maps the ring $K[f_1, \dots, f_N]$ into itself. Let $w_1, \dots, w_m$ be distinct complex numbers not lying among the poles of the $f_i$, such that*

$$f_i(w_\nu) \in K$$

*for all $i = 1, \dots, N$, and $\nu = 1, \dots, m$. Then $m \leqq 20\rho\,[K : \mathbf{Q}]$.*

As corollaries, we obtain the transcendence of several types of classical numbers.

Corollary 1 (Hermite-Lindemann). *Let $\alpha$ be an algebraic number $\neq 0$. Then $e^\alpha$ is transcendental.*

*Proof.* Suppose $e^\alpha$ is algebraic. Let $K$ be the field generated by $\alpha$ and $e^\alpha$ over $\mathbf{Q}$. Let $f, g$ be the two functions $f(t) = t$ and $g(t) = e^{\alpha t}$. Then the ring $K[f, g]$ is mapped into itself by $D = d/dt$, and $f, g$ are obviously algebraically independent. We note that $f, g$ also take on values in $K$ at all numbers $\alpha, 2\alpha, \dots, m\alpha$ for $m$ arbitrary large. This contradicts Theorem 1.

Corollary 2 (Gelfond-Schneider). *Let $\alpha, \beta$ be algebraic, $\alpha \neq 0, 1$ and $\beta$ irrational. Then $\alpha^\beta$ is transcendental.*

*Proof.* If $\alpha^\beta$ is algebraic, let $K$ be the field generated over $\mathbf{Q}$ by $\alpha$, $\beta$ and $\alpha^\beta$. We apply Theorem 1 to the two functions $e^t$ and $e^{\beta t}$, and to the set of integral multiples of $\log \alpha$ (for any determination of the logarithm) to obtain a contradiction as in Corollary 1.

The next two corollaries may be omitted by those who don't know about elliptic modular functions.

COROLLARY 3 (Schneider). *Let $\wp$ be a Weierstrass function, with algebraic $g_2$, $g_3$. If $\alpha$ is algebraic $\neq 0$, then $\wp(\alpha)$ is transcendental.*

*Proof.* The argument is similar to the proof of Corollary 1, using the functions $t$, $\wp(t)$, and the algebraic addition theorem for the $\wp$-function. If $\alpha$ happens to be a period, so a pole of $\wp$, then one may work with $\alpha/n$ for some positive integer $n$, and consider values of $\wp$ at integral multiples of $\alpha/n$ which are not periods, and at which $\wp$ is therefore defined.

COROLLARY 4 (Schneider). *Let $\tau$ be algebraic with* $\mathrm{Im}(\tau) > 0$, *and let $j$ be the elliptic modular function. If $j(\tau)$ is algebraic, then $\tau$ is a quadratic imaginary irrationality.*

*Proof.* One can find a Weierstrass $\wp$-function having algebraic $g_2$, $g_3$ and the corresponding value $j(\tau)$. Let $K$ be a number field containing $g_2, g_3, j(\tau)$ and $\tau$. Consider the two functions $f(t) = \wp(t)$ and $g(t) = \wp(\tau t)$. They cannot be algebraically independent over $K$. Indeed, we have $\tau = \omega_2/\omega_1$ (where $\omega_2$, $\omega_1$ are fundamental periods), and our functions take on algebraic values at all complex numbers $m\alpha$, where $\alpha = \omega_1/2$ and $m$ is an odd integer. (We use this trick to avoid poles of $\wp$.) It is a routine matter from the theory of elliptic functions to conclude that multiplication by $\tau$ maps the vector space generated over $\mathbf{Q}$ by the period lattice $(\omega_1, \omega_2)$ into itself, and consequently that $\tau$ is quadratic over $\mathbf{Q}$, and is not real.

The analysis needed to understand Corollaries 3 and 4 is covered in any text dealing with elliptic functions. We shall see later generalizations in higher dimensions.

We emphasize, however, that the proof of Theorem 1 in §2 and §3 is completely self-contained and elementary, using nothing more of the theory of complex variables than the maximum modulus principle. The main part of the proof is given in §3, and in §2 we shall prove some simple preliminary estimates involving derivatives.

The result of Theorem 1 is essentially best possible, without making further assumptions on the functions involved. For instance, $e^{Q(t)}$ takes on algebraic values at all zeros of a polynomial $Q(t)$ with algebraic coefficients, and there may well be a finite number of these.

## §2. *Estimation of derivatives*

Let

$$P(T_1, \ldots, T_N) = \sum \alpha_{(\nu)} M_{(\nu)}(T)$$

be a polynomial with complex coefficients $\alpha_{(\nu)}$, in $N$ variables $T_1, \ldots, T_N$, and let

$$Q(T_1, \ldots, T_N) = \sum \beta_{(\nu)} M_{(\nu)}(T)$$

be a polynomial with *real* coefficients $\geqq 0$. We say that $Q$ *dominates* $P$ if $|\alpha_{(\nu)}| \leqq \beta_{(\nu)}$ for all $(\nu)$, and we write $P \prec Q$. The following properties are then immediately verified.

If $Q_1$ dominates $P_1$ and $Q_2$ dominates $P_2$ then

$$P_1 + P_2 \prec Q_1 + Q_2 \qquad \text{and} \qquad P_1 P_2 \prec Q_1 Q_2.$$

Furthermore, if $D_i$ $(i = 1, \ldots, N)$ is the $i$-th derivative with respect to $T_i$, and if $Q$ dominates $P$, then

$$D_i P \prec D_i Q.$$

Let $P$ be a polynomial with complex coefficients as above. We let

$$|P| = \max |\alpha_{(\nu)}|$$

be the maximum of the absolute values of the coefficients. If the total degree of $P$ is $\leqq r$, then

$$P \prec |P|(1 + T_1 + \cdots + T_N)^r.$$

Since it is easy to take the derivative of the right-hand side with respect to any variable $T_i$, we have an easy way of estimating repeated formal derivatives of polynomials.

Let $K$ be a number field.

Let $S$ be a set of elements of $K$. We denote by $\|S\|$ the maximum of the absolute values of all conjugates of elements of $S$. By a *denominator* for $S$, we mean a positive integer which is a common denominator for all elements of $S$. If $P$ is a polynomial with coefficients in $K$, we let $\|P\|$ and a denominator for $P$ be defined in terms of the set of coefficients of the polynomial. We abbreviate denominator by den.

Lemma 1. *Let $K$ be a number field. Let $f_1, \ldots, f_N$ be functions, holomorphic on a neighborhood of a point $w \in \mathbf{C}$, and assume that $D = d/dt$ maps the ring $K[f_1, \ldots, f_N]$ into itself. Assume that $f_i(w) \in K$ for all $i$. Then there exists a number $C_1$ having the following property. Let $P(T_1, \ldots, T_N)$ be a polynomial with coefficients in $K$, of degree $\leqq r$.*

*If we set $f = P(f_1, \ldots, f_N)$, then we have, for all positive integers $k$,*

$$\|D^k f(w)\| \leqq \|P\| r^k k! C_1^{k+r}.$$

*Furthermore, there is a denominator for $D^k f(w)$ bounded by* $\operatorname{den}(P)C_1^{k+r}$.

*Proof.* There exist polynomials $P_i(T_1, \ldots, T_N)$ with coefficients in $K$ such that

$$Df_i = P_i(f_1, \ldots, f_N).$$

Let $\delta$ be the maximum of their degrees. There exists a unique derivation $\overline{D}$ on $K[T_1, \ldots, T_N]$ such that $\overline{D}T_i = P_i(T_1, \ldots, T_N)$. For any polynomial $P$ in $K[T]$, we have

$$\overline{D}(P(T_1, \ldots, T_N)) = \sum_{i=1}^{N} (D_i P)(T_1, \ldots, T_N) \cdot P_i(T_1, \ldots, T_N)$$

where $D_1, \ldots, D_N$ are the formal partial derivatives. The polynomial $P$ is dominated by

$$\|P\|(1 + T_1 + \cdots + T_N)^r,$$

and each $P_i$ is dominated by

$$\|P_i\|(1 + T_1 + \cdots + T_N)^\delta.$$

Thus $\overline{D}P$ is dominated by

$$\|P\| C_2 r(1 + T_1 + \cdots + T_N)^{r+\delta}.$$

Proceeding inductively, one sees that

$$\overline{D}^k P \prec \|P\| C_3^k r^k k!(1 + T_1 + \cdots + T_N)^{r+k\delta}.$$

Substituting values $f_i(w)$ for $T_i$, we have

$$D^k f(w) = \overline{D}^k P(f_1(w), \ldots, f_N(w)),$$

whence we obtain the desired bound on $D^k f(w)$. The second assertion concerning denominators is proved also by a trivial induction (even easier than the preceding one).

## §3. *The main proof*

We now come to the proof of Theorem 1. Let $f$, $g$ be two functions among $f_1, \ldots, f_N$ which are algebraically independent over $K$. Let $r$ be a positive integer divisible by $2m$. We shall let $r$ tend to infinity at the end of the proof.

Let

$$F = \sum_{i,j=1}^{r} a_{ij} f^i g^j$$

have coefficients $a_{ij}$ in $I_K$. Let $n = r^2/2m$. We can select the $a_{ij}$ not all equal to 0, and such that

$$D^k F(w_\nu) = 0$$

for $0 \leqq k < n$ and $\nu = 1, \ldots, m$. Indeed, we have to solve a system of $mn$ linear equations in $r^2 = 2mn$ unknowns. Note that

$$mn/(2mn - mn) = 1.$$

The size of the coefficients, and a denominator for them, are obtained as a result of Lemma 1. By Siegel's lemma, we see that we can take the $a_{ij}$ such that

$$\begin{aligned} \text{size}(a_{ij}) &\leqq n \log r + n \log n + (n + r)C_1 \\ &\leqq 2n \log n \end{aligned}$$

for $n$ sufficiently large.

Since $f$, $g$ are algebraically independent over $K$, our function $F$ is not identically zero. We let $s$ be the smallest integer such that all derivatives of $F$ up to order $s - 1$ vanish at all points $w_1, \ldots, w_m$ but such that $D^s F$ does not vanish at one of the $w_\nu$, say $w$. Then $s \geqq n$. Furthermore, using Lemma 1, we find at once

$$\text{size } D^s F(w) \leqq 5s \log s$$

for $n$ (and hence $s$) sufficiently large.

On the other hand, let $\theta$ be an entire function of order $\leqq \rho$, such that $\theta f$ and $\theta g$ are entire, and $\theta(w) \neq 0$. Then $\theta^{2r} F$ is entire. We consider the entire function

$$E(t) = \frac{\theta(t)^{2r} F(t)}{\prod_{\nu=1}^{m} (t - w_\nu)^s}.$$

Then $E(w)$ differs from $D^s F(w)$ by obvious factors, bounded by $C_4^s s!$ for some constant $C_4$. We use the maximum modulus principle to estimate $E$ on a circle of radius $R = s^{1/2\rho}$. On this circle, $|t - w_\nu|$ has approximately the same absolute value as $R$, and consequently,

$$|E(w)| \leqq |E|_R \leqq \frac{s^{3s} C_5^{2rR}}{R^{ms}}.$$

We therefore obtain the estimate

$$\log |D^s F(w)| \leqq 4s \log s - \frac{1}{2\rho} ms \log s.$$

Comparing this with the estimate for the size of $D^sF(w)$, we obtain at once $m \leqq 20\rho\,[K : \mathbf{Q}]$, as desired.

*Note:* We have made no particular effort to shrink the universal constant 20, and with a little care, one can obviously reduce it to a smaller constant. But this is essentially irrelevant, since the interesting thing is that in special cases, using additional structure on the functions $f_1, \ldots, f_N$, we can eliminate it altogether.

## §4. *The exponential map on groups*

Let $G$ be a group variety defined over a field $K$. Let $T_e(G)$ be its tangent space at the origin. We may view $T_e(G)$ as a vector space over any field containing $K$, but it has a basis defined over $K$, because one can define tangent vectors algebraically.

Over the complex numbers, we can identify the tangent space with $\mathbf{C}^d$ if $d = \dim G$. Thus the exponential map

$$\exp\colon \mathbf{C}^d \to G_{\mathbf{C}}$$

can be viewed as a complex analytic map from $\mathbf{C}^d$ into $G_{\mathbf{C}}$. Of course, we can perform any linear automorphism on $\mathbf{C}^d$ and still have an exponential map. However, if we normalize this exponential map so that $\mathbf{C}^d$ is extended from a vector space over $K$, then any further freedom we have in making linear automorphisms is restricted to such automorphisms over $K$. In terms of matrices, this means that the matrix of such a linear automorphism must have coefficients in $K$.

Finally, we observe that to normalize the exponential map as above is equivalent to saying that the partial derivatives $\partial/\partial t_1, \ldots, \partial/\partial t_d$ (of the $d$ complex variables $t_1, \ldots, t_d$) are defined over $K$, in other words map the field of rational functions $K(G)$ into itself.

Let $V$ be a variety (algebraic, irreducible) defined over a field $K$. If $P$ is a point of $V$ with affine coordinates $(x_1, \ldots, x_n)$ then we denote by $K(P)$ the field $K(x_1, \ldots, x_n)$. We say that $P$ is *rational* over $K$ if $K(P) = K$, *algebraic* over $K$ if $K(P)$ is algebraic over $K$, and *transcendental* over $K$ if $K(P)$ is not algebraic over $K$.

THEOREM 2. *Let $G$ be a group variety defined over the field of algebraic numbers* $\mathbf{K}$. *Let $T_e$ be its tangent space at the origin, with its structure of* $\mathbf{K}$*-vector space. Let $\alpha$ be a non-zero element of $T_e$ which is rational over* $\mathbf{K}$ *(i.e. algebraic), and let* $\exp = \exp_G$ *be the exponential map on $G$, normalized to have algebraic derivative at the origin. Assume that* $\exp(t\alpha)$ *is not a rational function of $t$. Then* $\exp(\alpha)$ *is transcendental over* $\mathbf{K}$.

*Proof.* Assume first that $G$ is a linear group variety, thereby admitting a global matrix representation over $K$. A tangent vector at the origin is simply a matrix $M$, and $\exp(M)$ is given by the series

$$\exp(M) = \sum M^{\mu}/\mu!\,.$$

If $B$ is an invertible matrix, then

$$\exp(B^{-1}MB) = B^{-1}\exp(M)B.$$

Assume that $B$ is rational over $\mathbf{K}$. We observe that $\exp(M)$ is rational over $\mathbf{K}$ if and only if $\exp(B^{-1}MB)$ is rational over $\mathbf{K}$. Any matrix $M$ rational over $\mathbf{K}$ can be conjugated into a matrix consisting of blocks

$$\begin{pmatrix} \alpha & & & \\ 0 & \alpha & & N \\ \vdots & & \ddots & \\ 0 & \cdots & & \alpha \end{pmatrix} = \alpha I + N,$$

where $\alpha$ is algebraic, and $N$ is a rational nilpotent matrix. Therefore $\exp(\alpha I + N)$ is equal to $e^{\alpha}I$ times a rational matrix, and is transcendental if and only if $\alpha \neq 0$. Consequently, in the linear case, Theorem 2 is equivalent with the transcendence of $e^{\alpha}$ for algebraic $\alpha \neq 0$, which followed from Theorem 1.

In general, $G$ contains a maximal linear subgroup $L$ such that $G/L$ is an abelian variety $A$. Let $\pi\colon G \to A$ be the canonical homomorphism. Then $\pi$ is an algebraic mapping. Let $\pi_*$ be the induced homomorphism on the tangent spaces at the origin. If $\pi_*\alpha = 0$, then $\alpha$ is contained in the Lie algebra of $L$, and the theorem reduces to the linear case. If $\pi_*\alpha \neq 0$, then it suffices to prove that

$$\pi \circ \exp_G(\alpha) = \exp_A \circ\, \pi_*(\alpha)$$

is transcendental on $A$, and our theorem is reduced to the case when $G$ is an abelian variety. We note that in that case, the exponential map is always a transcendental function. For example, when $\dim A = 1$, we can represent this map by the $\wp$-function, so that viewing $A$ in the projective plane, we have

$$\exp(t) = (1, \wp(t), \wp'(t)),$$

in terms of projective coordinates. Thus when $\dim A = 1$, our theorem was covered by Corollary 3 of Theorem 1. In this case, the normalization amounts to specifying that the invariants $g_2$, $g_3$ of $\wp$ must be algebraic.

For arbitrary abelian varieties, we have the representation

$$\Theta: \mathbf{C}^d \to A_{\mathbf{C}}$$

as a quotient of $\mathbf{C}^d$, by means of theta functions. A tangent vector at the origin is then represented by coordinates

$$\alpha = (\alpha_1, \ldots, \alpha_d),$$

and is algebraic if and only if all $\alpha_i$ are algebraic. Our normalization of the exponential map means that the partial derivatives $\partial/\partial t_i$ are defined over the field of algebraic numbers, and

$$\exp(t\alpha) = \Theta(t\alpha).$$

Thus in the case of abelian varieties, Theorem 2 can be formulated as follows.

THEOREM 3. *Let $A$ be an abelian variety of dimension $d$, defined over the field of algebraic numbers $\mathbf{K}$. Let*

$$\Theta: \mathbf{C}^d \to A_{\mathbf{C}}$$

*be the homomorphism given by the theta functions, inducing an isomorphism of the complex torus onto $A_{\mathbf{C}}$. Assume that the derivations*

$$\partial/\partial t_i \quad (i = 1, \ldots, d)$$

*are defined over $\mathbf{K}$. If $\alpha \in \mathbf{C}^d$ is a complex vector $\neq 0$ such that all $\alpha_i$ are algebraic, then $\Theta(\alpha)$ is transcendental over $\mathbf{K}$.*

*Proof.* All algebraic objects in Theorem 3 are in fact defined over a finitely generated extension of $\mathbf{Q}$, so that without loss of generality, we may replace $\mathbf{K}$ by a number field $K$ (finite over $\mathbf{Q}$). In particular, we may assume $\exp(\alpha)$ is rational over $K$, and so is $\exp(t\alpha)$ for $t = 1, 2, \ldots, m$. We take $m > 40\,[K : \mathbf{Q}]$ (so that we can apply Theorem 1 later) and let these $m$ points be $w_1, \ldots, w_m$. Let

$$\Theta = (\theta_0, \ldots, \theta_N)$$

be the projective coordinates of our map. We can find a suitable linear combination of $\theta_0, \ldots, \theta_N$ with coefficients in $K$ which does not vanish at any of our $m$ points, so that after a projective change of coordinates, we may assume that it is $\theta_0$. We let $f_j$ be the meromorphic function induced by $\theta_i/\theta_0$ on the straight line $t\alpha$ in $\mathbf{C}$, so that

$$f_j(t) = \theta_j/\theta_0(t\alpha)$$

for $j = 1, \ldots, N$. We note that each $\theta_j/\theta_0$ is an abelian function, holomorphic at our $m$ points. Since the partial derivative of a function is holomorphic at every point where the function is holomorphic, it follows that the ring

$$K[\theta_1/\theta_0, \ldots, \theta_N/\theta_0]$$

is mapped into itself by the partial derivatives $\partial/\partial t_i$ $(i = 1, \ldots, d)$, because of our assumption that these partial derivatives are defined over $K$. But then $D = d/dt$ maps the ring

$$K[f_1, \ldots, f_N]$$

into itself.

The exponential map $\Theta$ cannot be an algebraic function, and hence the transcendence degree of the field

$$K\big(t, f_1(t), \ldots, f_N(t)\big)$$

over $K$ is at least 2. We are now in a position to apply Theorem 1, with $\rho = 2$ to conclude the proof of Theorem 3, and therefore the proof of Theorem 2.

Corollary. *Notation being as in Theorem 3, the periods of $\Theta$ are transcendental.*

*Remark.* As is well known, if $\exp\colon \mathbf{C}^d \to G_{\mathbf{C}}$ is the exponential map, then its inverse mapping, which should also be called the log in general, is none other than the abelian integral of vector-differentials of the first kind on $G$. Thus if $P \in G_{\mathbf{C}}$, then $u = \log P \in \mathbf{C}^d$ means that $P = \exp(u)$. Different values of $\log P$ differ by periods of the exponential map. Varying $G$, e.g. taking for $G$ the generalized Jacobian varieties of a curve, one obtains the various abelian integrals (of all three kinds), and Theorem 2 says that (under the hypotheses stated there), if $P$ is an algebraic point on $G$, then $\log P$ is transcendental. However, $\log P$ is a vector, and similarly, if $\alpha$ is algebraic, $\alpha \neq 0$, then $\exp(\alpha)$ has $d$ local coordinates. In either case, one knows that at least one coordinate is transcendental. It becomes a difficult problem to determine whether *all* the coordinates (belonging to local uniformizing parameters) are transcendental.

Theorem 2 expresses on group varieties a generalization of the transcendence of $e^\alpha$ for $\alpha$ algebraic $\neq 0$. We can also express a generalization for the transcendence of $\alpha^\beta$.

Theorem 4. *Let $G$ be a group variety, defined over a number field, and assume that its exponential map can be represented by meromorphic functions of finite order. Let $\varphi\colon \mathbf{C} \to G_{\mathbf{C}}$ be a 1-parameter subgroup, normalized to have algebraic derivative at the origin. If there exists a point $u \in \mathbf{C}$,*

*$u \neq 0$ such that $\varphi(u)$ is an algebraic point of $G$, then $\varphi$ has algebraic dimension 1, and $\varphi(\mathbf{C})$ is a 1-dimensional group subvariety of $G$.*

*Proof.* This follows from Theorem 1 like Theorems 2 and 3.

We note that Theorem 4 applies to linear groups, abelian varieties, and products of these. The representation of the exponential map by meromorphic functions of finite order is not explicitly in the literature for arbitrary group varieties, except in special cases (e.g. the Weierstrass zeta function which can be used to parametrize certain 2-dimensional group varieties, and Severi's quasi-abelian functions). We note that one needs only to consider commutative group varieties, since the Zariski closure of the 1-parameter subgroup is commutative.

## Historical note

Hermite proved the transcendence of $e$, and several years later, Lindemann proved the transcendence of $e^{\alpha}$ for algebraic $\alpha \neq 0$ by an extension of the method of Hermite. This method involved constructing a function with high zeros, but was formally quite different from the one we have used, and we shall make comments on this again later.

The method we have used is essentially the method of Gelfond and Schneider, used to prove the transcendence of $\alpha^{\beta}$, properly axiomatized and formulated. The differential equation and the addition theorem for the exponential function are used by both, and Schneider had made a partial attempt at axiomatization in [28]. However, Schneider's formulation there was too special to be applied, say to abelian functions, and the arrangement of his proof is still comparatively complicated.

Theorem 2 had been conjectured by Cartier, and was proved in [16]. The transcendence of the periods of abelian functions is due to Schneider [27], who had to devise an argument in several variables to get them. Because of our more general theorem on arbitrary algebraic differential equations, we can handle this result still as one involving only a single variable.

Lindemann proved also a theorem on algebraic independence, namely, if $\alpha_1, \ldots, \alpha_m$ are algebraic, and linearly independent over $\mathbf{Q}$, then

$$e^{\alpha_1}, \ldots, e^{\alpha_m}$$

are algebraically independent. This will be proved by Siegel's method in Chapter VII. For the ordinary exponential function, Schanuel has made a very general conjecture, to the effect that *if $\alpha_1, \ldots, \alpha_m$ are complex numbers, linearly independent over $\mathbf{Q}$, then the transcendence degree of*

$$\alpha_1, \ldots, \alpha_m, e^{\alpha_1}, \ldots, e^{\alpha_m}$$

*is at least m.* From this statement, one would get most statements about algebraic independence of values of $e^t$ and $\log t$ which one feels to be true. For instance, it has been conjectured for a long time (by anybody who has looked at the subject) that if $\beta_1, \ldots, \beta_m$ are non-zero, multiplicatively independent algebraic numbers, then

$$\log \beta_1, \ldots, \log \beta_m$$

are algebraically independent. Similarly, one could deduce the algebraic independence of $e$ and $\pi$, by considering

$$1,\ 2\pi i,\ e,\ e^{2\pi i}.$$

I would also conjecture that $\pi$ cannot lie in the field obtained by starting with the algebraic numbers, adjoining values of the exponential function, taking algebraic closure, and iterating these two operations. It is an exercise to show that this follows from Schanuel's conjecture.

More generally, given a group variety $G$ defined over a number field $K$, one can ask for the transcendence degree of the point $(\alpha, \exp(\alpha))$, where $\alpha$ is a tangent vector at the origin, and exp is normalized as in Theorem 2. We shall discuss this question again in the next chapter, since it involves problems in several variables.

Theorem 2 and Theorem 3 give a reasonably satisfactory beginning for the theory of transcendental points on group varieties—the main problem is to determine transcendence degrees. However, there is an interesting direction where the transcendence question is still open. Let $V$ be a projective, non-singular curve (variety of dimension 1) defined over a number field, and of genus $\geqq 2$. Let $U$ be a disc, centered at the origin in $\mathbf{C}$. Let $\varphi: U \to V_{\mathbf{C}}$ be the universal covering map, i.e. essentially the universal uniformizing map of the Riemann surface $V_{\mathbf{C}}$, normalized so that the origin goes to an algebraic point, and the tangent map $\varphi'(0)$ is algebraic. One conjectures that if $P$ is an algebraic point of the disc, $\neq 0$, then $\varphi(P)$ is transcendental on $V$. As a side question, one can ask if the radius of the disc is transcendental. In this problem, as well as in Theorem 1, it is the normalization of the map which allows one to distinguish between algebraic and transcendental points.

When dealing with differential equations, the normalization can be taken to be algebraic initial conditions, and algebraic coefficients. We shall meet again such a situation in Chapter VII. However, not all classical maps satisfy a differential equation, and the study of automorphic functions from the point of view of transcendental numbers is barely begun.

We conclude by returning to Hermite's proof of the transcendence of $e$. This proof involves the continued rational fractions associated with the power series of $e^t$. It is by means of these fractions that Hermite constructs

the function having a zero of high order. This is quite an interesting direction, which relates the theory of transcendental numbers with the theory of diophantine approximations. So far, continued fractions have been developed mostly as a one variable theory, and hence are still not very good for generalizing Hermite's proof to proofs of algebraic independence for other types of functions beside $e^t$. There is no doubt, however, that eventually one will have to have such a generalization, which will give quantitative results concerning classical transcendental numbers. These are known as measures of transcendence, and we shall return in greater detail to this point at the end of Chapter VI and Chapter VII, where such measures of transcendence will be obtained for some special numbers.

Independently of transcendence problems, one can raise an interesting question of algebraic-analytic nature, namely given a 1-parameter subgroup of an abelian variety (say Zariski dense), is its intersection with a hyperplane section necessarily non-empty, and infinite unless this subgroup is algebraic? This question arises in connection with Theorem 4 of Chapter II, and Theorem 3 of the present chapter.

## CHAPTER IV

# Functions of Several Variables

### §1. *Partial differential equations*

We shall extend the main theorem of Chapter III to functions of several variables.

We recall that an entire function of $d$ variables $g(U) = g(u_1, \ldots, u_d)$ is said to be of order $\leqq \rho$ if there exists a constant $C > 0$ such that

$$|g(U)| \leqq C^{R^\rho}$$

on the domain $|u_i| \leqq R$ for all sufficiently large $R$. A meromorphic function will be said to be of order $\leqq \rho$ if it can be written as a quotient of two entire functions of order $\leqq \rho$. If $P$ is a point in $\mathbf{C}^d$ and $f$ a meromorphic function of order $\leqq \rho$ we say that $f$ is defined at $P$ if we can write $f = g/h$ where $g, h$ are entire of order $\leqq \rho$ and $h(P) \neq 0$.

THEOREM 1. *Let $K$ be a number field, $f_1, \ldots, f_N$ meromorphic functions of order $\leqq \rho$ in $d$ variables. Assume that the ring $K[f_1, \ldots, f_N]$ is mapped into itself by the partial derivatives $D_1, \ldots, D_d$, and that its transcendence degree over $K$ is $\geqq d + 1$. Let $S$ be a finite set of points in $\mathbf{C}^d$ at which all the functions $f_\mu$ are defined, and such that $f_\mu(P) \in K$ for all $\mu$ and all $P \in S$. Suppose in addition that after some change of coordinate system, the set $S$ becomes a product*

$$S_1 \times \cdots \times S_d,$$

*where each $S_i$ is a set of $m$ distinct complex numbers. Then*

$$m \leqq b\rho\,[K : \mathbf{Q}]$$

*with a constant $b$ depending only on $d$, and easily estimated.*

The proof of Theorem 1 will follow the pattern of the one dimensional proof. We begin by preliminary remarks on estimates of partial derivatives and integrals.

## §2. *Estimates of derivatives and integrals*

We have the analogue in several variables for the estimate of Chapter III, §2.

LEMMA 1. *Let $K$ be a number field. Let $f_1, \ldots, f_N$ be functions of $d$ complex variables, holomorphic at a point $A$ in $\mathbf{C}^d$, and such that the ring $K[f_1, \ldots, f_N]$ is mapped into itself by the partial derivatives $D_1, \ldots, D_d$. There exists a number $C_1$ having the following property. If $Q(T_1, \ldots, T_N)$ is a polynomial in $N$ variables with coefficients in $I_K$, of total degree $\leqq r$, and if*

$$D = D_1^{k_1} \cdots D_d^{k_d}$$

*is a differential operator of order $k = k_1 + \cdots + k_d$, then*

$$\|D(Q(f_1, \ldots, f_N))(A)\| \leqq \|Q\| r^k k! C_1^{k+r}.$$

*Furthermore, there is a denominator for $D(Q(f_1, \ldots, f_N))(A)$ bounded by* $\operatorname{den}(Q) C_1^{k+r}$.

*Proof.* The proof uses exactly the same kind of estimates as the proof for one variable, and the same kind of induction. We may therefore leave it as a simple exercise to the reader.

Next, we discuss some applications of Cauchy's formula. When dealing with functions of one variable, it was convenient to make one estimate using the maximum modulus principle. This argument does not extend to functions of several variables, but in the application we have in mind, we can use Cauchy's integral theorem instead. For clarity, we recall some facts in one variable. Let $S$ be a set of $m$ distinct complex numbers, and let $E(z)$ be an entire function of one variable $z$. Let $R$ be a positive number such that $|y| < R$ for all $y \in S$. For each $y \in S$, suppose given an integer $\lambda(y) \geqq 0$. Let

$$Q(z) = \prod_{y \in S} (z - y)^{\lambda(y)},$$

and for each $y \in S$, let

$$Q_y^*(z) = \frac{Q(z)}{(z - y)^{\lambda(y)}}.$$

Then $Q_y^*$ is defined and not zero at $y$. Let

$$\Delta_y = (d/dz)^{\lambda(y)-1}$$

if $\lambda(y) \geqq 1$, and 0 otherwise. Then

$$\int_{|z|=R} \frac{E(z)}{Q(z)}\, dz = \frac{1}{2\pi i} \sum_{y \in S} \frac{1}{(\lambda(y) - 1)!} \Delta_y(E/Q_y^*)(y).$$

If we deal with a function of several variables, then we can form a repeated integral over a similar quotient, and obtain an analogous formula. We shall need a special case.

LEMMA 2. *Let $E(z_1, \ldots, z_d)$ be an entire function. Let each one of $S_1, \ldots, S^d$ be a set of $m$ complex numbers. Let*

$$\Delta_i = \partial/\partial z_i.$$

*Let $\sigma$ be an integer $> 0$. Assume that for all differential operators $\Delta$ of order $\leqq \sigma$ and all points $Y \in S_1 \times \cdots \times S_d$ we have*

$$\Delta E(Y) = 0.$$

*Let $(\lambda) = (\lambda_1, \ldots, \lambda_d)$ be a vector of integers $\geqq 0$ such that*

$$\lambda_1 + \cdots + \lambda_d = \sigma + 1,$$

*and assume that*

$$\Delta^{(\lambda)} E(A) \neq 0$$

*for some $A$ in $S_1 \times \cdots \times S_d$. Let $S_i' = S_i - \{a_i\}$. Let*

$$Q(Z) = (z_1 - a_1)^{\lambda_1+1} \cdots (z_d - a_d)^{\lambda_d+1} \prod_i \prod_{y_i \in S_i'} (z_i - y_i)^{\lambda_i}.$$

*Let $R > |Y|$ for all $Y$. Then*

$$\int \cdots \int \frac{E(Z)}{Q(Z)} \, dZ = \frac{1}{\lambda!} \frac{1}{(2\pi i)^d} \frac{\Delta^{(\lambda)} E(A)}{Q_A^*(A)}$$

*where the repeated integral is taken over $|z_i| = R$, and*

$$Q_A^*(Z) = \frac{Q(Z)}{(z_1 - a_1)^{\lambda_1+1} \cdots (z_d - a_d)^{\lambda_d+1}}.$$

*Proof.* Taking the repeated integral and using the remarks we made above for one variable, we obtain a sum over vectors $Y$ involving differential operators. However, when evaluated at vectors $Y$, they will give only a zero contribution because of our assumption that $\Delta E(Y) = 0$, except for the term belonging to $A$, which is the only term for which a differential operator of order $\sigma + 1$ will occur. Up to powers of $2\pi i$ and factorials, this term is of type

$$\Delta^{(\lambda)}(E/Q_A^*)|_{Z=A}.$$

Our lemma now follows from the next remark.

LEMMA 3. *Let $E_1$, $E_2$ be two functions holomorphic at a point $A$. Assume that $\Delta E_1(A) = 0$ for all differential operators $\Delta$ of order $\leqq \sigma$, and let $\Delta^{(\lambda)}$ be an operator of order $\sigma + 1$. Then*

$$\Delta^{(\lambda)}(E_1E_2)(A) = \Delta^{(\lambda)}E_1(A) \cdot E_2(A).$$

*Proof.* Using repeatedly the rule for the derivative of a product, with respect to $\Delta_1, \ldots, \Delta_d$, we find that all terms will be 0 except the term which applies all the differential operations to $E_1$ and not to $E_2$. This yields what we wanted.

## §3. *The main proof*

We carry out the proper part of the proof of Theorem 1, and follow the same pattern as the proof of Theorem 1, Chapter III, taking into account the additional parameter $d$. Constants depending only on $d$ will be denoted by $b_1, b_2, \ldots$ Constants depending on $K$, $f_1, \ldots, f_N$ and $S$ will be denoted by $c_1, c_2, \ldots$ .

For any positive integer $n$, we have

$$b_1 n^d \leqq \binom{n+d}{d} \leqq b_2 n^d.$$

For convenience, we assume that $b_2$ is the $d$-th power of an integer.

Let $f_1, \ldots, f_N$ be our functions in $d$ variables $U = (u_1, \ldots, u_d)$. We let $D_i = \partial/\partial u_i$. Say $f_1, \ldots, f_{d+1}$ are algebraically independent over $K$. Let $r$ be a large integer, which will tend to infinity later, which is a $d$-th power of an integer, such that $r^{1/d}$ is divisible by $2mb_2$. Let

$$F = \sum a_{(i)} f_1^{i_1} \cdots f_{d+1}^{i_{d+1}}$$

have coefficients $a_{(i)}$ in $I_K$, the sum being taken for

$$0 \leqq i_1, \ldots, i_{d+1} < r.$$

We can select the coefficients $a_{(i)}$ not all equal to 0, and such that

$$DF(P) = 0$$

for all differential operators $D$ of order at most

$$n = \frac{r^{(d+1)/d}}{b_2^{1/d} 2m}$$

and all $P \in S$. We have to solve a system of linear equations in $r^{d+1}$ unknowns, and the number of equations is at most $b_2 n^d m^d \leqq r^{d+1}/2^d$. Using

Lemma 1, we see that the size of the coefficients, namely the size of the expressions

$$D(f_1^{i_1} \cdots f_{d+1}^{i_{d+1}})(P),$$

is bounded by $b_3 r \log r$ for $r$ sufficiently large (with respect to the data of the theorem). By Siegel's lemma, we can solve our linear equations with

$$\operatorname{size}(a_{(i)}) \leqq b_4 r \log r.$$

Since the functions $f_1, \ldots, f_{d+1}$ are algebraically independent over $K$, it follows that $F$ is not identically zero.

Let $\sigma$ be the smallest positive integer such that

$$DF(P) = 0$$

for all differential operators $D$ of order $\leqq \sigma$ and all $P \in S$. Then $\sigma \geqq n$ and there exists a $d$-tuple $(s) = (s_1, \ldots, s_d)$ such that

$$s_1 + \cdots + s_d = \sigma + 1,$$

and

$$D^{(s)}F(\overline{A}) \neq 0$$

for some $\overline{A}$ in $S$. By Lemma 1,

$$\operatorname{size} D^{(s)}F(\overline{A}) \leqq b_5 \sigma \log \sigma,$$

for $n$, and hence $\sigma$, sufficiently large.

We shall now estimate one conjugate of $D^{(s)}F(\overline{A})$ using Lemma 2.

Let $B$ be the $d \times d$ matrix giving the change of coordinates mentioned in the theorem, so that $U = ZB$, and $Z = (z_1, \ldots, z_d)$, both $U, Z$ being row vectors. Then

$$f_i(U) = g_i(Z)$$

and

$$F(U) = F(ZB) = G(Z).$$

We let $\Delta_i = \partial/\partial z_i$. Then there are linear combinations $L_1, \ldots, L_d$ of $\Delta_1, \ldots, \Delta_d$ with complex coefficients such that

$$D_i F(U) = L_i G(Z),$$

and hence

$$D^{(s)}F(U) = L^{(s)}G(Z).$$

We can write

$$L^{(s)} = L_1^{s_1} \cdots L_d^{s_d} = \sum \xi_{(\lambda)} \Delta_1^{\lambda_1} \cdots \Delta_d^{\lambda_d}$$

with complex coefficients $\xi_{(\lambda)}$ and $\lambda_1 + \cdots + \lambda_d = \sigma + 1$.

Let $A$ be the point in $S_1 \times \cdots \times S_d$ corresponding to $\overline{A}$ under the linear transformation. We then have

$$D^{(s)}F(\overline{A}) = L^{(s)}G(A) = \sum \xi_{(\lambda)} \Delta_1^{\lambda_1} \cdots \Delta_d^{\lambda_d} G(A).$$

We let $Y$ range over $S_1 \times \cdots \times S_d$ (the transform of $S$ under our linear transformation). Since no point of $Y$ lies in the poles of the functions $g_i$, we can find an entire function $\theta$ of order $\leqq \rho$ such that $\theta(A) \neq 0$ and $\theta g_i$ is entire of order $\leqq \rho$. We let

$$H(Z) = \theta(Z)^{(d+1)r}.$$

Then

$$E(Z) = G(Z)H(Z)$$

is an entire function, and $H(A) \neq 0$. Furthermore,

$$\Delta E(Y) = 0$$

for all differential operators $\Delta$ of order $\leqq \sigma$, and by Lemma 3,

$$\Delta^{(\lambda)} E(Y) = \Delta^{(\lambda)} G(Y)H(Y).$$

By Lemma 2, we obtain

$$\Delta^{(\lambda)} G(A) \leqq \lambda!(2\pi)^d \frac{|Q_A^*(A)|}{|H(A)|} \left| \int \cdots \int \frac{E(Z)}{Q(Z)}\, dZ \right|,$$

taking the repeated integral on the circles of radius

$$R = \sigma^{1/(d+1)\sigma}.$$

We can easily estimate the integral, and see that $1/Q(Z)$ gives essentially a contribution bounded by $1/R^{m\sigma}$ as soon as $\sigma$ is large enough with respect to the points $Y$. One can also estimate $E(Z)$ and one finds a bound of type

$$|E(Z)| \leqq c_1^{rR^\rho} \leqq c_2^\sigma$$

for $\sigma$ sufficiently large.

We therefore find

$$\log |\Delta^{(\lambda)} G(A)| \leqq 2\sigma \log \sigma - m\sigma \log R,$$

whence, estimating the $\xi_{(\lambda)}$, the number of terms in the sum expressing $L^{(s)}G(A)$, and $H(A), Q_A^*(A)$ in a trivial manner, we find

$$\log |D^{(s)}F(\overline{A})| = \log |L^{(s)}G(A)| \leqq 2\sigma \log \sigma - \frac{m\sigma \log \sigma}{(d+1)\rho}.$$

Comparing this estimate with the size of the element $D^{(s)}F(\overline{A})$ in $K$, we conclude that

$$m \leqq b\rho\,[K : \mathbf{Q}]$$

for some constant $b$ depending only on $d$, thereby proving Theorem 1.

## §4. *Applications*

We begin by a weak version in higher dimensions for an analogue of Theorem 4, Chapter II, §4.

THEOREM 2. *Let $G$ be either an abelian variety, or a linear group variety, defined over a number field. Let*

$$\varphi: \mathbf{C}^d \to G_{\mathbf{C}}$$

*be a $d$-parameter subgroup, normalized so that the derivative at the origin is algebraic. Let $\Gamma$ be a subgroup of $\mathbf{C}^d$ containing at least $d$ linearly independent points over $\mathbf{C}$. If $\varphi(\Gamma)$ is contained in the group of algebraic points of $G$, then $\varphi$ has algebraic dimension $d$, and $\varphi(\mathbf{C}^d)$ is an algebraic subgroup of dimension $d$.*

*Proof.* The assumption that the points of $\Gamma$ are linearly independent over the complex numbers allows us to change coordinates in such a way that they become a product. In the present case, we obtain $\infty^d$ points in $\mathbf{C}^d$ at which our functions take on algebraic values. In the case of abelian varieties, we dehomogenize the projective map at a product of $m$ such points with $m$ large, with respect to the new coordinate system. We can then apply Theorem 2 directly.

It would be very desirable to have a version of Theorem 2 in which the map $\varphi$ is not necessarily normalized with respect to the derivative, but which allows a subgroup with sufficiently many independent points over the rationals. Cf. the historical note.

Next we shall obtain an analogue to the transcendence statement concerning the modular function given in Chapter III.

Let $A$ be an abelian variety defined over the number field $K$. Let

$$\Theta: \mathbf{C}^d \to A_{\mathbf{C}}$$

be a representation as a quotient of complex $d$-space. We *assume* throughout that our representation is normalized so that the derivative of $\Theta$ at the origin is defined over $K$. Let $\Omega_1, \ldots, \Omega_{2d}$ be fundamental periods, and let $\Omega$ be the $d \times 2d$ matrix of periods, viewing the $\Omega_i$ as column vectors. Let $\mathbf{P}$ be a principal matrix such that $\Omega$ and $\mathbf{P}$ satisfy the Riemann relations. After making an *integral* change of coordinates, we can assume that

$\mathbf{P}$ has the usual canonical form

$$\mathbf{P} = \begin{pmatrix} 0 & \mathbf{D} \\ -\mathbf{D} & 0 \end{pmatrix}.$$

Then $\Omega = (W_1, W_2)$, where each $W_1, W_2$ is a $d \times d$ matrix, and $W_1$ is invertible. In the theory of moduli, one takes an integral matrix multiple of $W_1^{-1}W_2$ as the moduli point associated with the abelian variety in the Siegel upper half space $H_d$. Our next theorem is concerned however with $W_2W_1^{-1}$. The relation between $W_1^{-1}W_2$ and $W_2W_1^{-1}$ is not clear.

THEOREM 3. *Let $A$ be as above, and also $\Theta\colon \mathbf{C}^d \to A_{\mathbf{C}}$, as well as the period matrix $\Omega$, normalized so that the principal matrix has the usual canonical form. Let $T = W_2W_1^{-1}$. If $T$ is algebraic, then $T$ (viewed as a linear transformation) maps the period lattice tensored with $\mathbf{Q}$ into itself.*
*(When $d = 1$, then $T = \tau$, $W_1 = \omega_1$, $W_2 = \omega_2$ in the usual notation.)*

*Proof.* Let $\Phi(U) = \Theta(TU)$ and assume that $T$ is algebraic. After extending $K$ if necessary, we may assume that all components of $T$ lie in $K$. We view $U$ as a vertical vector for this proof. Then $\Omega_1, \ldots, \Omega_d$ are periods for $\Phi$ and $\Theta$, and generate a lattice such that for any point $P$ in this lattice, $\Phi(P)$ and $\Theta(P)$ are points of $A$ rational over $K$. A change of basis in $\mathbf{C}^d$ transforms lattice points into a set of points as in Theorem 1, namely

$$(j_1, \ldots, j_d), \qquad 0 \leqq j_i \leqq m.$$

Taking $m$ sufficiently large, we dehomogenize the projective maps $\Phi$ and $\Theta$ to obtain two rings of functions

$$K[f_1, \ldots, f_N] \quad \text{and} \quad K[g_1, \ldots, g_M]$$

giving corresponding maps into affine space and defined at the above points. Applying Theorem 1, we conclude that the product map

$$(\Phi, \Theta)\colon \mathbf{C}^d \to A_{\mathbf{C}} \times A_{\mathbf{C}}$$

has in fact dimension $d$, i.e. that $\Phi$ is algebraic over $\Theta$. For some integer $r$, it follows that $\Phi(rU)$ is rational over $\Theta(U)$, and hence that *all* periods of $\Theta(U)$ are also periods of $\Phi(rU)$. Hence $T$ maps the period lattice tensored with $\mathbf{Q}$ into itself, as was to be shown.

## *Historical note*

Schneider was the first to consider the transcendence question for functions of several variables in a special case [27], and as we have said earlier, proved that the periods of algebraic abelian integrals of the first and second kind are transcendental. He uses Cauchy's formula, and the technique of

repeated integrals is taken from his paper. The results of this chapter are taken from [17], and generalize Schneider's results.

He also makes a comment on the moduli point, and presumably had something like our Theorem 3 in mind ([27], page 113, top), but there is some confusion about the matter, in view of the fact that we find $W_2W_1^{-1}$ and not $W_1^{-1}W_2$.

Functions of several variables have of course divisors of zeros (not isolated points), and thus in investigating other types of sets $S$ on which such functions take on algebraic values, one meets genuine difficulties in putting natural conditions on $S$. For instance, the functions $z_1$, $z_2$, $e^{z_1 - z_2}$ take on algebraic values whenever $z_1$, $z_2$ are algebraic and $z_1 = z_2$. Furthermore, already in one variable, there may be certain collapsing in transcendence degree, since the functions

$$t, e^t, e^{t^2}, \ldots$$

have values in a field of transcendence degree 1 whenever $t$ is an integer.

In extending Theorems 1, 2, 3, one has the following possibilities.

First, a suggestion of Nagata, that the set $S$ in Theorem 1 may be described as contained in a hypersurface (algebraic). This would be a good way of eliminating the unnatural condition that $S$ be a product. (One would also need to bound the degree of the hypersurface.)

Second, even though we know that the vector periods of abelian integrals are transcendental, this means only that at least *one* component is transcendental. It is therefore desirable to have stronger theorems, giving the transcendence of each component. In this direction, the Riemann relations give significant examples of degeneracy, as follows. Let $C_1, \ldots, C_m$ be elements of complex $d$-space, which are linearly independent over $\mathbf{Q}$. Let $A$ be an abelian variety of dimension $d$ defined over a number field $K$, and assume that its ring of endomorphisms is trivial. In particular, $A$ is simple. Let

$$\Phi_m : \mathbf{C}^{md} \to A^m_{\mathbf{C}}$$

be the mapping given by $\Phi_m(t) = (\Theta(tC_1), \ldots, \Theta(tC_m))$. Thus $\Phi_m$ is the exponential map passing through the vector $(C_1, \ldots, C_m)$. *We contend that this map has algebraic dimension md*, i.e. *that it is Zariski-dense in* $A^m$. We prove this by induction on $m$. For $m = 1$, it follows from the simplicity of $A$. Assume it for $m$. If the dimension of the map with $m + 1$ factors is not $d(m + 1)$, then $\Phi_{m+1}(t)$ is contained in an abelian subvariety of $A^{m+1}$ whose projection on the first $m$ factors is $A^m$, and which is algebraic over $A^m$. There exists an integer $r \neq 0$ such that

$$\Theta(rtC_{m+1})$$

is rational over $\Phi_m(t)$, i.e.

$$\big(\Theta(tC_1), \ldots, \Theta(tC_m), \Theta(rtC_{m+1})\big)$$

lies in an abelian subvariety $B$ of $A^{m+1}$ which is isomorphic to $A^m$ under projection, and whose projection on the last factor is $A$. Hence there exists a homomorphism of $A^m$ onto $A$, and since $A$ has only trivial endomorphisms, there exist integers $r_1, \ldots, r_m$ such that

$$\Theta(trC_{m+1}) = r_1\Theta(tC_1) + \cdots + r_m\Theta(tC_m).$$

This implies that

$$rC_{m+1} = r_1C_1 + \cdots + r_mC_m,$$

a contradiction.

We apply the above to the $2d$ periods $\Omega_1, \ldots, \Omega_{2d}$. The dimension of the map $\Phi_{2d}$ is then $2d^2$, but the period $(\Omega_1, \ldots, \Omega_{2d})$ is mapped onto 0 by $\Phi_{2d}$. According to the Riemann relations, the transcendence degree of the period matrix cannot be $2d^2$.

(On the other hand, observe that the Riemann relations do not give a counterexample for the transcendence of the components of a basic period vector.)

One may raise the question whether such degeneracy can occur in the case of a simple abelian variety, possibly under the additional assumption that the ring of endomorphisms is trivial. In other words, suppose that $A$ is a simple abelian variety. Let $\alpha$ be a non-zero algebraic tangent vector (i.e. an element of $\mathbf{C}^d$ which is algebraic). Then one may ask whether the transcendence degree of the point $\big(\alpha, \Theta(\alpha)\big)$ in $\mathbf{C}^d \times A_{\mathbf{C}}$ is $\geqq d$.

It would be desirable to have a general conjecture also for products. For instance, if $A$ has dimension 1, and our mapping is represented by the Weierstrass function with algebraic $g_2, g_3$, let $\omega_1, \omega_2$ be two fundamental periods. If $A$ has no complex multiplication, then it is reasonable to expect that $\omega_1, \omega_2$ are algebraically independent. One would see this by looking at the representation of $A \times A$ as a quotient of $\mathbf{C}^2$.

For the period matrix itself, Grothendieck has made a very interesting conjecture concerning its relations, and his conjecture applies to a general situation, as follows. Let $V$ be a projective, non-singular variety defined over the rational numbers. One can define the cohomology of $V$ with rational coefficients in two ways. First, by means of differential forms (de Rham), purely algebraically, thereby obtaining a vector space $H_{\text{diff}}(V, \mathbf{Q})$ over $\mathbf{Q}$. Secondly, one can take the singular cohomology $H_{\text{sing}}(V, \mathbf{Q})$ with rational coefficients, i.e. the singular cohomology of the complex manifold $V_{\mathbf{C}}$. Let us select a basis for each of these vector spaces

over $\mathbf{Q}$, and let us tensor these spaces over $\mathbf{C}$. Then there is a unique (period) matrix $\Omega$ with complex coefficients which transforms one basis into the other. Any algebraic cycle on $V$ or the products of $V$ with itself will give rise to a polynomial relation with rational coefficients among the coefficients of this matrix. Grothendieck's conjecture is that the ideal generated by these relations is an ideal of definition for the period matrix.

We observe that one can also consider group varieties which are not complete, and that the transcendental parametrizations of these are the inverse maps of abelian integrals of various kinds. The transcendence and algebraic independence of values of abelian integrals thus appears as a question on the inverse mapping of the general exponential map on group varieties, i.e. as the algebraic independence of logarithms in an extended sense. This would be the case for instance for the integral

$$\int_0^1 \frac{1}{1+t^3}\,dt = \frac{1}{3}\left(\log 2 + \frac{\pi}{\sqrt{3}}\right)$$

which appears at the end of Siegel's book. Of course, this integral involves only the algebraic independence of ordinary logarithms ($2\pi i = \log 1$ and $\log 2$). More generally, one can consider an intermediate case related to the Weierstrass zeta function ($\zeta' = -\wp$), which, together with the $\wp$-function, parametrizes the group variety associated with integrals of the second kind by the map $\mathbf{C} \times \mathbf{C} \to G_{\mathbf{C}}$ given by

$$(t, u) \mapsto \big(1, \wp(t), \wp'(t), u - \zeta(t)\big).$$

This map has periods $(\omega_1, \eta_1)$ and $(\omega_2, \eta_2)$ (classical notation), and the Legendre relation

$$\eta_1\omega_2 - \eta_2\omega_1 = 2\pi i$$

is nothing but the Riemann relation for that case.

Schanuel's conjecture mentioned in Chapter III asserts that such exceptions do not occur for the ordinary exponential function. It would be interesting to determine if there are other examples of degeneracy for the transcendence degree of the point $\big(\alpha, \exp(\alpha)\big)$ whenever the analytic curve $\exp(t\alpha)$ has the correct dimension, other than those given by period relations.

For abelian varieties, the transcendence degree is affected not only by the Riemann relations, but also by the existence of non-trivial endomorphisms, or of algebraic cycles on the variety and its products. In any case, one is also led in this way to investigate the transcendence degree of points $\big(\alpha, \varphi(\alpha)\big)$, where

$$\varphi\colon H_d \times \mathbf{C}^d \to V_{\mathbf{C}}$$

is the uniformization of the variety of moduli $V$ by means of the Siegel upper half space $H_d$, and complex $d$-space.

In any case, the Riemann relations show that one cannot prove the algebraic independence of logarithms of multiplicatively independent algebraic numbers by "general" statements about entire functions. One must make use of the special properties of $e^t$.

Third, as we have already mentioned, one would like a higher dimensional version of Theorems 3 and 4, Chapter I, §3 and §4, corresponding to Theorem 2 of this chapter, but without assuming any normalization on the parametrization of the $d$-parameter subgroup. One then meets two difficulties in trying to extend the one-dimensional proof. First, a purely analytic difficulty which consists in formulating a suitable interpolation estimate. The use of Cauchy's formula does not seem to generalize without some new idea. In any case, it is obviously necessary to assume that the linear combinations of the log vectors of algebraic points with integer coefficients are essentially equidistributed. This leads to the second difficulty, much deeper than the first, namely to prove that such combinations satisfy this equidistribution property. This is a problem which by definition belongs to the theory of diophantine approximations, and at present is very much beyond any result which has ever been obtained in this theory.

CHAPTER V

# Finitely Generated Values

## §1. The inductive technique

Up to now, we have considered the problem of proving the transcendence of values of certain functions over the field of rational numbers. However, in cases of numbers like

$$e + \pi \qquad \text{or} \qquad e + \log 2$$

we find that the number is a sum of values of *two* different types of functions (namely exp and log), so that the natural approach to the transcendence of such values is to prove that say $e$, $\pi$ are algebraically independent. So far, such a result has eluded available techniques, but we shall point out a good possibility in this direction, and obtain partial results. We work out a special case first. The general case will involve only additional technical complications.

Let $x$ be a complex number, transcendental over $\mathbf{Q}$. For any polynomial $P(x) \in \mathbf{Z}[x]$ with integer coefficients, we define

$$\operatorname{size} P(x) = \max(\deg P, \log |P|).$$

Let $\tau$ be a number $\geqq 2$. We shall say that $x$ is of *transcendence type* $\leqq \tau$ if for all polynomials with integral coefficients, we have

$$-(\operatorname{size} P(x))^{\tau} \ll \log |P(x)|.$$

This is a natural extension of the formula stated in Chapter I for algebraic numbers (in that case, $\tau = 1$). From Dirichlet's box principle, one sees that one can never take $\tau < 2$. (Our definition is adjusted to the method for which it will be useful. Since our results are only partial, a refinement of the method will require a corresponding refinement of the definition, to prove more refined results.)

To estimate products of polynomials, we note that

$$\operatorname{size}(P_1 \cdots P_n) \leqq 3 \sum_{i=1}^{n} \operatorname{size} P_i.$$

This is trivially proved, because $P_1 \cdots P_n$ is dominated by

$$|P_1| \cdots |P_n|(1 + T)^{d_1 + \cdots + d_n},$$

where $d_i = \deg P_i$, and hence

$$|P_1 \cdots P_n| \leqq |P_1| \cdots |P_n| 2^{d_1 + \cdots + d_n},$$

whence our estimate follows. (The absolute value around $P_i$ means of course the maximum of the coefficients, and is to be distinguished from $|P_i(x)|$, taken in the complex numbers.)

We illustrate the generalization of the technique of Chapter I by a simple statement.

THEOREM 1. *Let $\beta_1, \ldots, \beta_d$ be complex numbers, linearly independent over $\mathbf{Q}$. Let $z_1, \ldots, z_m$ be complex numbers linearly independent over $\mathbf{Q}$. Let $x$ be a complex number of type $\leqq \tau$. If $dm \geqq \tau(m + d)$ then not all values*

$$e^{\beta_i z_\nu} \qquad (i = 1, \ldots, d \text{ and } \nu = 1, \ldots, m)$$

*can lie in the field $\mathbf{Q}(x)$.*

*Proof.* For simplicity, assume that all values above are polynomials, in $\mathbf{Z}[x]$. (In the general case, one has merely to clear denominators throughout the proof.) Let $n$ be a large integer, and let $r$ be approximately equal to $n^{m/d}$, up to a constant factor. We form the function

$$F(t) = \sum a_{(i)} x^{i_{d+1}} e^{i_1 \beta_1 t} \cdots e^{i_d \beta_d t}$$

with

$$1 \leqq i_1, \ldots, i_d \leqq r \qquad \text{and} \qquad 1 \leqq i_{d+1} \leqq nr,$$

with integer coefficients $a_{(i)}$, not all zero, such that $F$ has a zero at all numbers

$$k \cdot z = k_1 z_1 + \cdots + k_m z_m$$

with

$$1 \leqq k_\nu \leqq n.$$

We have to solve linear equations, and for each $(k)$, substituting $k \cdot z$ for $t$ in $F(t)$, we obtain a polynomial in $x$ with integer coefficients. Trivial estimates show that the degrees of these polynomials are $\ll rn$. Thus the number of linear equations, bounded by the number of $(k)$ times a bound for the degrees, satisfies

$$\text{number of equations} \ll rn^{m+1}.$$

On the other hand,

$$\text{number of variables} = nr^{d+1}.$$

Thus our choice of $r$ guarantees that the exponent in Siegel's lemma is constant. Simple estimates show that the coefficients of our linear equations satisfy

$$\text{size of coefficients} \ll nr.$$

Hence we can solve our equations with integers $a_{(i)}$ not all 0, of size $\ll nr$. We then let $s$ be the largest integer such that $F(k \cdot z) = 0$ for $1 \leqq k_\nu \leqq s$, and let $w = k \cdot z$ with some $k_\nu = s + 1$ be such that $F(w) \neq 0$. Then we estimate

$$F(w) = \frac{F(t)}{\prod(t - k \cdot z)} \prod (w - k \cdot z)\Big|_{t=w},$$

using the maximum principle on the circle of radius $R = s^{1+1/d}$. We find that the quotient of the products pulls towards zero with a strength at least

$$\frac{1}{d} s^m \log s,$$

whereas the numerator pulls to infinity with a strength at most

$$s^{m/d} sR = s^{m/d+2+1/d}.$$

This is safely smaller than the strength of the denominator, and we obtain

$$\log |F(w)| \ll -s^m \log s.$$

On the other hand, simple estimates show that

$$\text{size } F(w) \ll sr \ll s^{m/d+1},$$

the size being of course the size of the polynomial in $x$. By assumption, we must have

$$-(s^{m/d+1})^\tau \ll -s^m \log s,$$

which contradicts our assumption on $m$ and $d$.

We observe that in Theorem 1, we must always have $d > \tau$, and that if we do, then we can choose $m$ sufficiently large so that the condition $dm \geqq \tau(m + d)$ is satisfied. For instance, if $\tau = 2$, we can take $d = 3$ and $m = 6$.

We shall see in the next sections that one may deal with any finite extension of a field $\mathbf{Q}(x)$ and obtain a similar result. This is a matter only of technique, as is the generalization to finitely generated extensions $\mathbf{Q}(x_1, \ldots, x_q, y)$ with algebraically independent $x_1, \ldots, x_q$, and $y$ algebraic over $\mathbf{Q}(x)$. One then has to assume a transcendence type for the elements $(x)$, and the pattern of the proof is the same.

If instead of $\mathbf{Q}(x)$ we deal with a finite extension of $\mathbf{Q}(x)$, then we see that Theorem 1 yields a result on algebraic independence, namely one of the values

$$e^{\beta_i z_\nu}$$

is algebraically independent from $x$. Thus a theorem like Theorem 1 yields an inductive procedure for algebraic independence results. However, it is very imprecise, even more so than the analogous Theorem 1 of Chapter II, where we already mentioned the desirability of shrinking 3 to 2. It seems that to go deeper, one must refine the notion of transcendence type, and load the induction hypotheses much more than we have done, in addition to changing something in the structure of the proof. A discussion of this problem is best postponed until the end of Chapter VI, in the light of Feldman's results.

## §2. *Transcendence types*

Let $K$ be a subfield of the complex numbers, finitely generated over $\mathbf{Q}$. Suppose that we write

$$K = \mathbf{Q}(x_1, \ldots, x_q, y),$$

where $x_1, \ldots, x_q$ are algebraically independent, $y$ is algebraic over $\mathbf{Q}(x)$, and is integral over $\mathbf{Z}[x_1, \ldots, x_q] = \mathbf{Z}[x]$. One can always find such generators, and a set $(x, y)$ satisfying these conditions will be called a *proper set of generators.*

With respect to such a proper set $(x, y)$, we define the *size* of an element of $K$ as follows. We do it first for polynomials, i.e. elements of $\mathbf{Z}[x]$. If

$$\alpha = \sum c_{(i)} x_1^{i_1} \cdots x_q^{i_q} = P(x)$$

with coefficients $c_{(i)} \in \mathbf{Z}$, we let

$$\operatorname{size}(\alpha) = \max(\deg P, \log |c_{(i)}|).$$

Thus we let the size be the maximum of the degree of $P$ and the logs of the absolute values of the coefficients. We shall also call it the size of $P$, and write it $\sigma(P)$.

Next, let $\alpha \in \mathbf{Z}[x, y]$. Then we can write

$$\alpha = P_0(x) + P_1(x)y + \cdots + P_{N-1}(x)y^{N-1},$$

where $P_0(x), \ldots, P_{N-1}(x) \in \mathbf{Z}[x]$, if $N$ is the degree of $y$ over $\mathbf{Q}(x)$. We define

$$\operatorname{size}(\alpha) = \max \text{ sizes } P_0(x), \ldots, P_{N-1}(x).$$

Finally, if $\gamma$ is an element of $K$, we say

$$\operatorname{size}(\gamma) \leqq B$$

if $\gamma$ can be written as a quotient $\gamma = \alpha/\beta$ with $\alpha \in \mathbf{Z}[x, y]$, $\beta \in \mathbf{Z}[x]$, such that $\operatorname{size}(\alpha)$, $\operatorname{size}(\beta) \leqq B$. Thus the size of an element of $K$ is defined in a natural way, similar to our definition of the size of an algebraic number.

LEMMA 1. *Let the notation be as above. There exists a number $c > 0$ (depending on $x, y$) such that, if $\alpha_1, \dots, \alpha_m$ are elements of $K$, then*

$$\operatorname{size}(\alpha_1 \cdots \alpha_m) \leqq c(\operatorname{size}(\alpha_1) + \cdots + \operatorname{size}(\alpha_m)),$$
$$\operatorname{size}(\alpha_1 + \cdots + \alpha_m) \leqq c(\operatorname{size}(\alpha_1) + \cdots + \operatorname{size}(\alpha_m)).$$

*Proof.* We first prove the first statement for polynomials in $\mathbf{Z}[x]$. If $P_1(x), \dots, P_m(x)$ are such polynomials, of degrees $d_1, \dots, d_m$ respectively, then their product is dominated by

$$P_1(T) \cdots P_m(T) \prec |P_1| \cdots |P_m|(1 + T_1 + \cdots + T_q)^{d_1 + \cdots + d_m},$$

and if we expand out the power of $(1 + T_1 + \cdots + T_q)$ on the right, we see that the coefficients are bounded by

$$c^{d_1 + \cdots + d_m}$$

for some constant $c$. Taking the log yields what we want.

Next, suppose that our elements $\alpha_1, \dots, \alpha_m$ are written

$$\alpha_k = P_{k1}(x) + \cdots + P_{k,N-1}(x)y^{N-1}$$

for $k = 1, \dots, m$. Introduce a new variable $T_{q+1}$, and let

$$A_k = P_{k1}(T) + \cdots + P_{k,N-1}(T)T_{q+1}^{N-1}.$$

Then the product $A_1(T) \cdots A_m(T)$ is dominated in the same way that we argued previously for polynomials, by a sum

$$|A_1| \cdots |A_m| c^{d_1 + \cdots + d_m} \sum T_1^{i_1} \cdots T_q^{i_q} T_{q+1}^{i_{q+1}},$$

and $i_1, \dots, i_{q+1} \leqq d_1 + \cdots + d_m + N$. Let $g(Y)$ be the irreducible polynomial of $y$ over $\mathbf{Z}[x]$, with leading coefficient 1 since we assumed $y$ integral over $\mathbf{Z}[x]$. For each $i$ we can write

$$Y^i = g(Y)h_i(Y) + r_i(Y)$$

with polynomial $h_i$, $r_i$ in $\mathbf{Z}[Y]$. A simple induction, using long division, shows that the size of the coefficients of the remainder term $r_i(Y)$ is

bounded by $cd$ for some constant $c$. We substitute each such remainder term $r_i(T_{q+1})$ for the corresponding power of $T^i_{q+1}$ in our sum above, and thus see again that if we substitute $(x_1, \ldots, x_q)$ for $(T_1, \ldots, T_q)$ and $y$ for $T_{q+1}$, we obtain a bound for the size of the desired type.

The second statement concerning the size of a sum is trivially proved from the first, putting all terms in the sum over a common denominator.

It will be convenient to use the resultant in the next lemma. If

$$f(Y) = a_n Y^n + \cdots + a_0$$
$$g(Y) = b_m Y^m + \cdots + b_0$$

with indeterminate coefficients $a_i$, $b_j$, then

$$\mathrm{R}(f, g) = \begin{vmatrix} a_n & \cdots\cdots & a_0 & & \\ & a_n & \cdots\cdots & a_0 & \\ & & \cdots\cdots\cdots & & \\ & & a_n & \cdots\cdots & a_0 \\ b_m & \cdots\cdots & b_0 & & \\ & b_m & \cdots\cdots & b_0 & \\ & & \cdots\cdots\cdots & & \\ & & b_m & \cdots\cdots & b_0 \end{vmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} m \\ \left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\} n \end{matrix}$$

and this resultant is also equal to

$$R(f, g) = \begin{vmatrix} a_n & \cdots\cdots & a_0 & & & Y^{m-1}f \\ & a_n & \cdots\cdots & a_0 & & Y^{m-2}f \\ & & \cdots\cdots\cdots & & & \vdots \\ & & a_n & \cdots\cdots & a_0 & Yf \\ & & & a_n & \cdots\cdots\ a_1 & f \\ b_m & \cdots\cdots & b_0 & & & Y^{n-1}g \\ & b_m & \cdots\cdots & b_0 & & Y^{n-2}g \\ & & \cdots\cdots\cdots & & & \vdots \\ & & b_m & \cdots\cdots & b_0 & Yg \\ & & & b_m & \cdots\cdots\ b_1 & g \end{vmatrix}$$

If $w$ is a complex number, and the coefficients $a_i$, $b_j$ of $f$, $g$ are taken to be complex numbers, with $a_n b_m \neq 0$, then we obtain, substituting $w$ for $Y$ in the second determinant,

$$\boxed{|R(f, g)| \leqq (1 + |w|)^{m+n}\big(|f(w)| + |g(w)|\big)|f|^m|g|^n(m + n)^{m+n}.}$$

LEMMA 2. *There exists a constant $c > 0$ such that for all $\alpha \neq 0$ in $K$ we have* $\text{size}(\alpha) \leqq c \cdot \text{size}(1/\alpha)$.

*Proof.* Without loss of generality, we may assume that

$$\alpha = P_1(x) + \cdots + P_{N-1}(x)y^{N-1} = f(y),$$

where $P_1, \ldots, P_{N-1}$ are polynomials with integer coefficients. Let $g(Y)$ be the irreducible polynomial of $y$ over $\mathbf{Z}[x]$. Since $f(y) \neq 0$, the resultant of $g$ and $f$, which is an element of $\mathbf{Z}[x]$, is not zero. Let $P(x)$ be this resultant. Then

$$P(x) = R(f, g) = Af + Bg,$$

where $A, B$ are polynomials, elements of $\mathbf{Z}[x][Y]$. From the expression of the resultant as a determinant, and from Lemma 1, it is trivial to see that the size of the coefficients of $A$ and $B$ (in $\mathbf{Z}[x]$) is bounded by $c \cdot \text{size}(\alpha)$ for some constant $c$. Furthermore, the degree of $A$ in $y$ is bounded by a constant depending on $N$. Substituting $y$ for $Y$ in our expression for the resultant, we find that

$$P(x) = A(y)f(y),$$

and thus $\alpha^{-1} = A(y)/P(x)$. From the preceding remarks, we find that $\text{size}(\alpha^{-1}) \ll \text{size}(\alpha)$.

We shall say that $K$ has *transcendence type* $\leqq \tau$ (for some number $\tau \geqq 2$) if there exists a proper set of generators $(x, y)$ such that, with respect to this set, for every non-zero element $\alpha \in K$, we have

$$-(\text{size } \alpha)^{\tau} \ll \log |\alpha|.$$

LEMMA 3. *Assume that $\mathbf{Q}(x)$ has transcendence type $\leqq \tau$, with respect to $(x)$. If $y$ is a generator of $K$ over $\mathbf{Q}(x)$, and $y$ is integral over $\mathbf{Z}[x]$, then $K$ has transcendence type $\leqq \tau$ with respect to $(x, y)$.*

*Proof.* Write each $\alpha \in K$, $\alpha \neq 0$ in the form

$$\alpha(x, y) = \frac{1}{P(x)}\,(P_0(x) + \cdots + P_{N-1}(x)y^{N-1}) = \frac{1}{P(x)}\,f(y),$$

where $P, P_0, \ldots, P_{N-1}$ have size $\leqq \sigma_\alpha$, and

$$\log |\alpha| \ll -c_\alpha \sigma_\alpha^{\tau}$$

for some positive number $c_\alpha$. If the assertion of our lemma is false, then we can find a sequence of numbers $\alpha$ for which $c_\alpha$ tends to infinity. Multiplying $\alpha$ by $P(x)$, and observing that $|P(x)| \leqq c^{\sigma_\alpha}$ for some constant $c$, we see that, without loss of generality, we may assume that $P(x) = 1$. Let $g$ be the irreducible polynomial of $\alpha$ over $\mathbf{Z}[x]$, and let $Q_\alpha(x) = R(f, g)$

be the resultant of $f$ and $g$. Using the estimate for the resultant, resulting from the determinant expression, we find

$$|Q_\alpha(x)| \leqq |\alpha| c^{\sigma_\alpha}$$

for some constant $c$. Again, the determinant expression of the resultant shows that

$$\text{size}\, Q_\alpha(x) \ll \text{size}\, \alpha.$$

Hence we obtain

$$\log |Q_\alpha(x)| \ll -c_\alpha \sigma_\alpha^\tau \ll -c_\alpha (\text{size}\, Q_\alpha(x))^\tau.$$

This contradicts the fact that the field $\mathbf{Q}(x)$ has transcendence type $\leqq \tau$ with respect to $(x)$, and proves our lemma.

The preceding lemmas show that working with a finite extension of the pure transcendental field $\mathbf{Q}(x)$, from the present point of view, is essentially no harder than working with $\mathbf{Q}(x)$ itself, involving only routine polynomial technique. In particular, the result of the preceding section would hold just as well for such a field $K$. We shall formulate it more generally in the next section.

## §3. *Algebraic independence*

We have all the tools to extend the definitions and results of Chapter II, §2 to the case of functions taking values in a finitely generated field $K$, of transcendence degree $q$ and transcendence type $\leqq \tau$. We always assume that $\tau \geqq q + 1 \geqq 2$.

Let

$$S = \bigcup S_n$$

be again a set expressed as a union of subsets $S_n$, with $S_n \subset S_{n+1}$ for all $n = 1, 2, \ldots$ We assume again that $|z| \leqq Cn$ for all $z \in S_n$. A meromorphic function $f$ defined on $S$, of order $\leqq \rho$, with values in $K$, is said to be of arithmetic order $\leqq \rho$ on $S$ if there is a constant $C \geqq 1$ such that the following conditions are satisfied:

AO 1. *For all $n$ and $z \in S_n$ we have* size $f(z) \leqq Cn^\rho$.

AO 2. *There is an entire function $h$ of order $\leqq \rho$, such that $hf$ is entire, $h$ has no zero in $S$, and for all $n$, $z \in S_n$,*

$$\log |1/h(z)| \leqq Cn^\rho.$$

Observe that the statements of AO 1 and AO 2 are identical with those of Chapter II.

THEOREM 2. *Let $f_1, \ldots, f_d$ be meromorphic functions, of order $\leqq \rho$, defined on the set $S$ as above, with values in $K$, and of arithmetic order $\leqq \rho$ on $S$. Let $m$ be a number such that $(m + d)\tau \leqq m(d + q - 1)$. Assume that* card$(S_n) \gg\ll n^{m\rho}$, *for $n \to \infty$. Then $f_1, \ldots, f_d$ are algebraically dependent over $K$.*

*Proof.* Suppose that $f_1, \ldots, f_d$ are independent over $K$. Let $n$ be as usual a large integer which will tend to infinity. Let $r$ be approximately equal to

$$n^{(m-q+1)\rho/(d+q-1)}.$$

We form the function

$$F = \sum a_{(i)} x_1^{i_{d+1}} \cdots x_q^{i_{d+q}} f_1^{i_1} \cdots f_d^{i_d},$$

with

$$1 \leqq i_1, \ldots, i_d \leqq r \qquad \text{and} \qquad 1 \leqq i_{d+1}, \ldots, i_{d+q} \leqq rn^\rho.$$

We require that $F(z) = 0$ for all $z \in S_n$. This amounts to solving the usual linear equations, setting the coefficient of each monomial in $(x)$ equal to 0. We have

$$\text{number of unknowns} = r^d(rn^\rho)^q.$$

The degrees in $(x)$ of the elements $F(z)$ with $z \in S_n$ is $\ll rn^\rho$. Hence

$$\text{number of equations} \ll n^{m\rho} rn^\rho.$$

Our assumption on $r$ means that the number of unknowns is approximately equal to the number of equations, and thus that the exponent in Siegel's lemma is bounded by a constant. The size of each $F(z)$, with $z \in S_n$, is also bounded by $rn^\rho$, up to a constant factor. Hence we can solve for the coefficients $a_{(i)}$ so that they have size $\ll rn^\rho$.

Let $s$ be the largest integer such that $F(z) = 0$ for all $z \in S_s$, and let $w \in S_{s+1}$ be such that $F(w) \neq 0$. Then $s \geqq n$ and

$$\text{size } F(w) \ll rs^\rho.$$

We estimate $|F(w)|$ as usual, from the expression

$$F(w) = \frac{h(t)^{dr}F(t)}{h(w)^{dr}\prod(t - z)} \prod (w - z) \Big|_{t=w},$$

where $h$ is a function satisfying AO 2 for each one of $f_1, \ldots, f_d$. We take the circle of radius $R = s^{1+1/dq}$ (any $s^{1+\epsilon}$ would do). Then the entire function $h(t)^{dr}F(t)$ on the circle of radius $R$ is bounded by

$$\log |h^{dr}F|_R \ll rR^\rho \ll rs^{(1+1/dq)\rho} \ll s^{m\rho}.$$

On the other hand, the quotient of the two products in the estimate for $F(w)$ pulls towards zero, and in fact

$$\log \max_{|t|=R} \left| \frac{\prod(w-z)}{\prod(t-z)} \right| \ll -s^{m\rho} \log s.$$

The power $h(w)^{dr}$ does not affect anything by hypothesis. Hence essentially only the product counts in the estimate for $|F(w)|$, and we get

$$\log |F(w)| \ll -s^{m\rho} \log s.$$

Now we get a contradiction of the inequality

$$-(\text{size } F(w))^{\tau} \ll \log |F(w)|,$$

using our hypothesis relating $m$, $d$ and $\tau$ (the hypothesis was in fact made up so that at this point in the proof, we get the desired contradiction). This proves Theorem 2.

Corollary 1. *Assume that we are given $d > \tau$, and let $m$ be such that $m(d - \tau) \geqq d\tau$. Then with the other hypotheses of Theorem 2, the functions $f_1, \ldots, f_d$ are algebraically dependent over $K$.*

Corollary 2. *Let $\mathbf{Q}(x)$ be a purely transcendental extension of $\mathbf{Q}$, of transcendence type $\leqq \tau$ for some integer $\tau \geqq 2$. Let $G$ be the general linear group, of some dimension, and $\mathbf{K}$ the algebraic closure of $\mathbf{Q}(x)$. Let $\varphi \colon \mathbf{C} \to G_{\mathbf{C}}$ be a 1-parameter subgroup of $G$, of algebraic dimension $d$. Let $\Gamma$ be a subgroup of $\mathbf{C}$, containing at least $m$ elements linearly independent over $\mathbf{Q}$, such that $\varphi(\Gamma) \subset G_{\mathbf{K}}$. If $m \geqq d\tau$, then $d \leqq \tau$.*

*Proof.* As in Chapter I, the estimates for the exponential series are easily carried out, to show that Corollary 1 applies.

*Remark.* One also wants a statement similar to Corollary 2 for abelian varieties. Everything goes through as in Chapter II, except that we used the Néron-Tate form to verify that condition AO 1 is satisfied, in the analogous situation. Here, the necessary verification has not yet been made. However we observe that the Néron-Tate form was a much more powerful tool than actually needed. All we need are upper estimates on the size of a sum of points on an abelian variety. This type of consideration belongs in a separate treatment of the problem on abelian varieties, which is purely algebraic. Note that it would admit the possibility that the abelian variety is defined over a finitely generated field, of finite transcendence type!

Aside from the algebraic independence of $e$, $\pi$ or $e$, $\log 2$, or $\log 2$, $\log 3$, the next simplest case to treat is $\wp(\alpha_1)$, $\wp(\alpha_2)$, where $\alpha_1$, $\alpha_2$ are linearly independent over $\mathbf{Q}$, and $\wp$ is a Weierstrass function with algebraic $g_2$, $g_3$

such that the corresponding elliptic curve has only trivial endomorphisms. It may be that this would come from pushing further the ideas of Feldman [4], some of which bear a vague analogy with Siegel's, except that Feldman deals with the algebraic differential equation of the $\wp$-function. Feldman's success with a certain inversion associated with the $\wp$-function, and this analogy, indicate that one may reach a proof of the similar result on abelian varieties, by a deepening of the method, still modelled on the usual pattern of constructing the function $F$ with a lot of zeros.

## Historical note

Gelfond proved a statement analogous to our Theorem 1 [13], but with the following substantial differences. To begin with, he does not make the assumption on the transcendence type of the field of values, and he uses the differential equation of the exponential function. Furthermore, he uses in an incidental way a measure of irrationality for exponents $\beta_i$. Finally, he uses a very special feature about the exponential function, which makes it very unclear how to extend his proof to more general functions. On the other hand, in place of the transcendence type, he has the following theorem, valid for *all* numbers. (For a proof, cf. [21].)

THEOREM. *Let $x$ be a complex number. Let $\sigma$ be a strictly monotone increasing real function tending to infinity, and assume that there is a number $a_0 > 1$ such that $\sigma(N+1) < a_0\sigma(N)$ for all integers $N > N_0$. Assume that for each integer $N > N_0$ there exists a non-zero polynomial $F_N$ with integer coefficients, such that*

$$|F_N(x)| < e^{-C\sigma^2(N)}$$

*where $C$ is a sufficiently large constant, and*

$$\max(\deg F_N, \log |F_N|) \leqq \sigma(N).$$

*Then $x$ is algebraic.*

The difficulty about applying this theorem is that one needs very many polynomials $F_N$ having small values at $x$. This means that, in an analogous situation to our Theorem 1, he must have a way of showing that the integer $s$ is not too large, and in fact is about the same order of magnitude as the integer $n$. It is here that he uses the above-mentioned special properties of the function $e^t$. In addition, other technical complications arise in the course of the proof. Thus in spite of much greater complications, he still ends up only with partial results, as we do here.

The method I use in this chapter is much simpler than Gelfond's method, and has the additional advantage of showing clearly the inductive rela-

tionship between transcendence types and the possibility of proving results of algebraic independence. For instance, we could state Theorem 2 for fields of transcendence degree $> 1$, whereas Gelfond could not state a similar theorem. Of course, the problem remains of proving that certain numbers have definite types, and to solve this problem, one expects a higher order of complication, of the nature encountered by Feldman in his papers (cf. Chapter VI). It then becomes clear that one must refine the notion of transcendence type, in a manner to be discussed at the end of Chapter VI.

*A transcendence measure* for a number $x$ is any function $g$ of two variables such that for all polynomials $P$ with integer coefficients, $P \neq 0$, of degree $\leqq d$ and height $\leqq h$ one has

$$\log |P(x)| \geqq g(d, h).$$

The problem of determining the best possible transcendence measures for classical numbers is by definition a problem in diophantine approximations, which is thereby shown to be inseparable from the theory of transcendental numbers.

As we said above, the result of this chapter is quite weak, and is only intended to show in a simple case the first example of an inductive procedure which could eventually be used to obtain best possible results. It will then be necessary to refine the method of proof. This can come from the following directions:

(1) Use linear inequalities, rather than the linear equations, which are extremely wasteful in the number of variables. Cf. Chapter VI, §3.

(2) Use an improved transcendence measure. It seems that the whole inductive procedure is set up in such a way that only an extremely refined inductive assumption can lead to the proper result. We shall discuss in Chapter VI what such assumptions could be like.

(3) Even using such assumptions, there is still something missing in the present structure of the proof, even using an algebraic differential equation, or for concreteness the functions $t$, $e^t$. As far as I can tell, making any and all of these assumptions, it is still not possible with the present structure of the proof to derive a contradiction leading to the best possible conjecturable results.

# CHAPTER VI

# Transcendence Measures

## §1. *The Liouville estimate*

We shall reformulate a somewhat more general result than the fundamental inequality of Chapter I, §1, to deal with algebraic numbers whose degree is not necessarily fixed.

If $P$ is a polynomial with integer coefficients, we define its *height* $h(P)$ to be

$$h(P) = \log |P|,$$

i.e. it is the log of the maximum of the absolute values of its coefficients. Similarly, if $\xi$ is an algebraic number, and $P$ is its irreducible polynomial over $\mathbf{Z}$, then we define its height,

$$h(\xi) = h(P).$$

We define its *absolute size* $\sigma(\xi)$ to be

$$\sigma(\xi) = \max(\deg \xi, h(\xi)).$$

We shall now reformulate the fundamental inequality of Chapter I, §1 to deal with the absolute size of an algebraic number.

LEMMA 1. *Let*

$$P(X) = a_d(X - \alpha_1) \cdots (X - \alpha_d), \qquad a_d \neq 0$$

*be a polynomial with complex coefficients. Then*

$$|a_d| \prod_{i=1}^{d} \max(1, |\alpha_i|) \leqq 2^d |P|.$$

*Proof.* Dividing both sides by $|a_d|$, we may assume without loss of generality that $a_d = 1$. We now use induction on the number of indices $i$ such that $|\alpha_i| > 2$. If $|\alpha_i| \leqq 2$ for all $i$, our assertion is obvious. Suppose now that

$$P(X) = g(X)(X - \alpha)$$

with $|\alpha| > 2$, and suppose that our assertion is true for

$$g(X) = X^d + b_{d-1}X^{d-1} + \cdots + b_0.$$

We have

$$P(X) = X^{d+1} + (b_{d-1} - \alpha)X^d + (b_{d-2} - \alpha b_{d-1})X^{d-1} + \cdots + (-\alpha)b_0.$$

We can assume $|g| = |b_i| \geqq |b_{i-1}|$ for some $i$, $0 \leqq i \leqq d$ (with the convention $b_d = 1$, $b_{-1} = 0$). Then

$$\begin{aligned} |P| \geqq |\alpha b_i - b_{i-1}| &\geqq |\alpha|\,|b_i| - |b_{i-1}| \\ &\geqq |\alpha|\,|b_i| - |b_i| = (|\alpha| - 1)|b_i| \\ &\geqq \tfrac{1}{2}|\alpha|\,|b_i| = \tfrac{1}{2}|\alpha|\,|g|, \end{aligned}$$

and our lemma is now obvious, since $|\alpha| > 2$.

*Remark.* If in Lemma 1 we take a polynomial whose coefficients lie in a field with a non-archimedean valuation, then we have a similar inequality *without the factor* $2^d$. This is nothing but the Gauss lemma for valuations, and is trivially proved. In fact, we have the *equality*

$$|a_d|_p \prod_{i=1}^{d} \max(1, |\alpha_i|_p) = |P|_p$$

if $|\ |_p$ denotes a $p$-adic valuation.

THEOREM 1. *Let $\xi_1, \ldots, \xi_m$ be algebraic numbers, of degrees $d_1, \ldots, d_m$ and heights $h_1, \ldots, h_m$ respectively. Let*

$$d = [\mathbf{Q}(\xi_1, \ldots, \xi_m) : \mathbf{Q}].$$

*Let $P$ be a polynomial in $m$ variables $X_1, \ldots, X_m$, with integer coefficients, of degree $N_i$ in $X_i$. If $P(\xi_1, \ldots, \xi_m) \neq 0$, then*

$$-d\left[h(P) + \sum_{i=1}^{m} \frac{N_i h_i}{d_i} + 2\sum_{i=1}^{m} N_i\right] \leqq \log|P(\xi_1, \ldots, \xi_m)|.$$

*Proof.* Let $P_i$ be the irreducible polynomial of $\xi_i$ over $\mathbf{Z}$, and let $a_i$ be its leading coefficient. Assume $P(\xi) \neq 0$. Let $j = 1, \ldots, d$ be indices for the embeddings of $\mathbf{Q}(\xi)$ into the complex numbers. Since $P(X_1, \ldots, X_m)$ is dominated by

$$|P|(1 + X_1)^{N_1} \cdots (1 + X_m)^{N_m},$$

it follows that for each $j$,

$$|P(\xi_1^{(j)}, \dots, \xi_m^{(j)})| \leqq |P|(1 + |\xi_1^{(j)}|)^{N_1} \cdots (1 + |\xi_m^{(j)}|)^{N_m}.$$

Taking the product over $j = 1, \dots, d$ and using Lemma 1, applied to the irreducible polynomials of $\xi_1, \dots, \xi_m$ respectively, we find

$$\prod_{j=1}^{d} |P(\xi)^{(j)}| \leqq |P(\xi)|\,|P|^d \prod_{i=1}^{m} [4^{d_i} a_i^{-1} |P_i|]^{dN_i/d_i}.$$

On the other hand, we let $|\ |_p$ denote the absolute value on the algebraic closure of the $p$-adic field $\mathbf{Q}_p$, and apply Lemma 1 in the $p$-adic case, using $(j)$ to denote $p$-adic conjugates. We then obtain the similar estimate,

$$\begin{aligned}\prod_{j=1}^{d} |P(\xi)^{(j)}|_p &\leqq \prod_{i=1}^{m} \prod_{j=1}^{d} \sup(1, |\xi_i^{(j)}|_p)^{N_i} \\ &\leqq \prod_{i=1}^{m} |a_i^{-1} P_i|_p^{dN_i/d_i}.\end{aligned}$$

Since $P_i$ has integral coefficients, $|P_i|_p \leqq 1$, we can delete it from our $p$-adic estimate. Taking the product over all $p$, and over the ordinary absolute value, and using the product formula, we obtain

$$1 \leqq |P(\xi)|\,|P|^d \prod_{i=1}^{m} [4^{d_i} |P_i|]^{dN_i/d_i}.$$

Taking the log gives the estimate of the theorem.

We shall call the estimate of Theorem 1 a *Liouville estimate.*

COROLLARY. *Let $P$ be a polynomial with integer coefficients, and $\xi$ an algebraic number. If $P(\xi) \neq 0$, then*

$$-[\deg(\xi) h(P) + (\deg P) h(\xi) + 2 \deg P] \leqq \log |P(\xi)|.$$

## §2. *Polynomial and algebraic approximations*

Our next task is to investigate the relationship between polynomial and algebraic approximations.

LEMMA 2. *Let $P$ be a polynomial of degree $d$ with integer coefficients. Then every factor $Q$ of $P$ over $\mathbf{Z}$ satisfies the inequality*

$$|Q| \leqq 4^d |P|.$$

*Proof.* Any factor $Q$ of $P$ over $\mathbf{Z}$ can be written in the form

$$Q(X) = b_{d_1} \prod_{\nu=1}^{d_1} (X - \alpha_{i_\nu}),$$

where $\alpha_{i_\nu}$ ($\nu = 1, \ldots, d_1$) are roots of $P$, and $b_{d_1}$ is an integer dividing $a_d$, the leading coefficient of $P$. Then

$$|Q| \leqq |b_{d_1}| 4^d \left| \frac{1}{a_d} P \right| \leqq 4^d |P|,$$

as was to be shown.

COROLLARY. *Notation as in the lemma, we have* $\sigma(Q) \leqq 3\sigma(P)$, *where* $\sigma$ *is the size,* $\sigma(P) = \max(\deg P, h(P))$.

*Proof.* Obvious.

The next lemma shows that given a polynomial with integer coefficients taking a small value at a number $w$, we can find an irreducible factor which also has a small value at $w$.

LEMMA 3. *Let $P(X)$ be a polynomial with integer coefficients, of degree $d \geqq 1$, and let $\lambda$ be a positive number. If*

$$\log |P(w)| \leqq -\lambda d \sigma(P),$$

*then there exists an irreducible factor $P_1$ of $P$ over $\mathbf{Z}$ of degree $d_1$ such that*

$$\log |P_1(w)| \leqq -\tfrac{1}{3}\lambda d_1 \sigma(P_1).$$

*Proof.* Factor $P$ into a product of irreducible polynomials,

$$P = P_1 \cdots P_s.$$

Suppose that our assertion is false. Say

$$\log |P_i(w)| > -\tfrac{1}{3}\lambda d_i \sigma(P_i)$$

for $i = 1, \ldots, s$. Taking the sum, and using the Corollary of Lemma 2, we find

$$\begin{aligned} \log |P(w)| &> -\tfrac{1}{3}\lambda\big(d_1\sigma(P_1) + \cdots + d_s\sigma(P_s)\big) \\ &> -\lambda\sigma(P)(d_1 + \cdots + d_s) \\ &> -\lambda d\sigma(P), \end{aligned}$$

a contradiction which proves the lemma.

LEMMA 4. *Let*

$$P(X) = a_d X^d + \cdots + a_0$$

*be a polynomial with integer coefficients, $a_d \neq 0$, and without multiple roots. Let $w$ be a complex number, and let $\xi_1, \ldots, \xi_d$ be the roots of $P$. Then*

$$\min_i |w - \xi_i| \leqq |P(w)| e^{5d(\log d + h(P) + 4)}.$$

*Proof.* Let $\xi_1, \ldots, \xi_d$ be so ordered that

$$|w - \xi_1| \leqq \cdots \leqq |w - \xi_d|.$$

Then for all $i \neq 1$, we have

$$|w - \xi_i| \geqq \tfrac{1}{2}|\xi_1 - \xi_i|.$$

Otherwise,

$$|w - \xi_i| < \tfrac{1}{2}|\xi_1 - w| + \tfrac{1}{2}|w - \xi_i|,$$

which is impossible. Now we have

$$|P(w)| = |a_d|\,|w - \xi_1| \cdots |w - \xi_d| \geqq |w - \xi_1| 2^{-d+1} |P'(\xi_1)|.$$

We must therefore estimate $|P'(\xi_1)|$ from below. Note that $|\xi_1| \leqq d|P|$ trivially. Using the fundamental estimate for the resultant, we get

$$1 \leqq |R(P, P')| \leqq (1 + d|P|)^{2d} |P'(\xi_1)|\,|P|^{d-1} |P'|^d (2d)^{2d},$$

whence the desired estimate follows at once.

*Remark.* The estimate of the lemma is quite sharp in its dependence on $d$. Except for the constant 5 it is also sharp in its dependence on $h$. Variations of the proof can be given to eliminate the constant 5, as a coefficient of $h(P)$, at the cost of making the dependence on $d$ somewhat worse.

THEOREM 2. *Let $w$ be a transcendental number, and let $\lambda \geqq 1$ be a positive function of two real variables. Then the following conditions are equivalent:*

TM 1. *For all algebraic numbers $\xi$ of degree $\leqq d$ and absolute size $\leqq \sigma$, we have*

$$\log |w - \xi| \gg -\lambda(d, \sigma)\, d\sigma.$$

TM 2. *For all polynomials $P$ of degree $\leqq d$, with integer coefficients, and size $\leqq \sigma$, we have*

$$\log |P(w)| \gg -\lambda(d, \sigma)\, d\sigma.$$

*Proof.* Assume TM 1. To prove TM 2, we may assume by Lemma 2 that $P$ is irreducible over $\mathbf{Z}$. By Lemma 4, we find

$$\log \min |w - \xi_i| \ll \log |P(w)| + d\sigma.$$

Using TM 1 to get an inequality on the left, we see that $\log |P(w)|$ satisfies the desired inequality. Conversely, assume TM 2. To prove TM 1, we need only a more trivial inequality than that of Lemma 4, namely if $P$ is the irreducible polynomial of $\xi$ over $\mathbf{Z}$, then

$$|P(w)| \leqq |a_d|\,|w - \xi_1| \cdots |w - \xi_d|,$$

and since $|\xi_i| \leqq d|P|$ for all $i$, we obtain trivially

$$\log |P(w)| \ll \log |a_d| + \log |w - \xi| + d\sigma.$$

Applying TM 2 yields TM 1.

## §3. *Linear inequalities*

In order to make a function have small values at certain points, we shall need a substitute for the Siegel lemma, allowing us to solve linear inequalities instead of linear equations. This yields finer estimates than Siegel's lemma.

LEMMA 5. *Let*

$$\begin{aligned} L_1(X) &= \alpha_{11}x_1 + \cdots + \alpha_{1n}x_n \\ &\;\;\vdots \\ L_r(X) &= \alpha_{r1}x_1 + \cdots + \alpha_{rn}x_n \end{aligned}$$

*be a system of linear forms with complex coefficients $\alpha_{ij}$, and $n > 2r$. Let $A \geqq 1$ and $|\alpha_{ij}| \leqq A$ for all $i, j$. Let $B$ be a number $\geqq 1$. Then there exists a non-trivial integral solution $X$ of the inequalities*

$$|L_i(X)| \leqq \frac{1}{B} \quad \text{and} \quad |X| \leqq 2(8nAB)^{2r/(n-2r)}$$

*for all $i$.*

*Proof.* Assume first that the coefficients are real numbers, and only that $n > r$. We shall prove that we can satisfy the inequalities

$$|L_i(X)| \leqq \frac{1}{B} \quad \text{and} \quad |X| \leqq 2(4nAB)^{r/(n-r)}.$$

For any number $C \geqq 1$, the linear map $L\colon \mathbf{Z}^n \to \mathbf{R}^r$ given by our linear forms maps $\mathbf{Z}^n(C)$ into $\mathbf{R}^r(nAC)$. Cut the interval

$$-nAC \leqq t \leqq nAC$$

into $[4nACB]$ segments of equal length. Note that

$$[4nACB] \geqq 2nACB.$$

Then each small segment has length

$$\frac{2nAC}{[4nACB]} \leqq \frac{2nAC}{2nACB} \leqq \frac{1}{B}.$$

Then $\mathbf{R}^r(nAC)$ is decomposed into $[4nACB]^r$ small cubes of sides $\leqq 1/B$. In $\mathbf{Z}^n(C)$ we have at least $C^n$ integral vectors. If

$$C^n > [4nACB]^r,$$

then there exist two distinct vectors $Y, Y' \in \mathbf{Z}^n(C)$ such that $L(Y)$ and $L(Y')$ lie in the same small cube. For this, it suffices that

$$C > (4nAB)^{r/(n-r)}.$$

Then $X = Y - Y'$ satisfies our requirements.

Now in the complex case, for each linear form with complex coefficients, we write down two linear forms, the real and imaginary parts, separately. This yields twice as many equations, whence the needed assumption $n > 2r$. Furthermore, if $L'$ is the real part of $L$, and $L''$ its imaginary part, we solve

$$|L'(X)| \leqq 1/2B \qquad \text{and} \qquad |L''(X)| \leqq 1/2B,$$

so that we must replace $B$ by $2B$ in our real result to obtain the desired bound in the complex case.

## §4. *Interpolation estimates*

In the proof of this chapter, we shall not do as before, construct a function with many zeros, but rather, we construct a function having very small values at many points. This means that the maximum principle cannot be used any more, and that we shall use an interpolation method to estimate the function, obtained from an application of Cauchy's theorem. We state it as a separate lemma.

LEMMA 6. *Let $E$ be an entire function, and let*

$$Q(t) = [(t - z_1) \cdots (t - z_m)]^l$$

*be a polynomial with distinct roots $z_\nu$ ($\nu = 1, \ldots, m$) of multiplicity $l$. For each $\nu = 1, \ldots, m$ we let*

$$Q_\nu^*(t) = [(t - z_1) \cdots \widehat{(t - z_\nu)} \cdots (t - z_m)]^l,$$

*i.e. we omit the factor* $(t - z_\nu)^l$. *Let* $R$ *be a positive number such that* $|z_\nu| < R$ *for all* $\nu$, *and let* $z$ *be a number inside this circle, unequal to any* $z_\nu$. *Let* $\Gamma$ *be the circle of radius* $R$. *Then*

$$\frac{E(z)}{Q(z)} = \frac{1}{2\pi i}\int_\Gamma \frac{E(\zeta)}{Q(\zeta)}\frac{d\zeta}{\zeta - z}$$
$$- \frac{1}{2\pi i}\sum_{\nu=1}^{m}\sum_{k=1}^{l-1}\frac{1}{k!}D^kE(z_\nu)\int_{\Gamma_\nu}\frac{1}{Q_\nu^*(\zeta)(\zeta - z)}\frac{d\zeta}{(\zeta - z_\nu)^{l-k}},$$

*where* $\Gamma_\nu$ *is a circle around* $z_\nu$, *not containing* $z$, *and not containing any other* $z_\mu$.

*Proof.* We start with the formula given for one variable in Chapter IV, §2, and then expand the multiple derivative $D^{l-1}(EG)$ as a sum

$$\sum_{k=0}^{l-1}\binom{l-1}{k}D^kE\cdot D^{l-1-k}G,$$

where $G$ is an obvious function. We then express each $D^{l-1-k}G(z_\nu)$ by the usual Cauchy integral, and the desired formula comes out trivially.

From our expression, we see that if $D^kE(z_\nu)$ is small, then $E(z)$ will also be small, i.e. we have obtained the technical equivalent of the estimates which we would have if $E$ had zeros of multiplicity $l$ at all points $z_1, \ldots, z_m$. Thus $E(z)$ is bounded essentially in terms of $1/Q(\zeta)$, and $D^kE(z_\nu)$. The other expressions entering in our formula play a secondary role in practice. We shall put this formula in a form which is best adapted for the applications we have in mind, also giving an estimate for derivatives of $E$ (easily achieved using Cauchy's formula).

LEMMA 7. *Let the hypotheses be as in Lemma* 6. *Let* $2 < R_1 < R$. *Let* $z'$ *be a complex number, distinct from* $z_1, \ldots, z_m$, *and assume that* $z', z_1, \ldots, z_m$ *lie inside a circle of radius* $R_1/2$. *Let* $\delta$ *be less than the minimum of the distances of any pair of distinct numbers among*

$$z', z_1, \ldots, z_m,$$

*and also* $0 < \delta < 1$. *Then for any integer* $r \geqq 0$ *we have*

$$|D^rE(z')| \leqq \frac{2^r r!}{R_1^r}[(C_1R_1/R)^{ml}|E|_R + ml(1/\delta)^{ml}\max_{k,\nu}|D^kE(z_\nu)|]$$

*where* $C_1$ *is an absolute constant.*

*Proof.* We have

$$|D^rE(z')| \leqq |D^rE|_{R_1}.$$

We use the preceding lemma to estimate $|D^r E|_{R_1}$, and begin by estimating $|E|_{R_1}$, taking $z$ on the circle of radius $R_1$ in Lemma 6. We estimate $Q(z)$ on the circle of radius $R_1$, and $Q(\zeta)$ on the circle of radius $R$. Then for some universal constant $C_1$,

$$|E|_{R_1} \leqq C_1^{ml}(R_1/R)^{ml}|E|_R + mlC_1^{ml}(1/\delta)^{ml} \max_{k,\nu} |D^k E(z_\nu)|.$$

To get the $r$-th derivative, we use Cauchy's formula,

$$D^r E(z') = \frac{r!}{2\pi i} \int_{|t|=R_1} \frac{E(t)\,dt}{(t - z')^{r+1}}$$

and apply our estimate for $|E|_{R_1}$ to prove our lemma.

In the applications, we must make two estimates to apply Lemma 7, corresponding to the two terms in the sum.

The first will be small because of $(R_1/R)^{ml}$, and will be determined by this expression. It is then necessary to verify that $r!$ and $C_1^{ml}$, as well as $|E|_R$, do not tend to infinity faster than $(R_1/R)^{ml}$ tends to 0.

The second will be small because of $\max_{k,\nu} |D^k E(z_\nu)|$, and will be determined by these derivatives. It is then necessary to show that $r!$, $C_1^{ml}$, and $(1/\delta)^{ml}$ again do not tend to infinity faster than the absolute value of the derivatives involved.

## §5. *A determinant*

Let $f$, $g$ be two meromorphic functions of a complex variable. Let $\varphi_{ij}$ be complex numbers, and let us form the function

$$F = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \varphi_{ij} f^i g^j.$$

Then

$$D^k F = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \varphi_{ij} D^k(f^i g^j),$$

and

$$D^k(f^i g^j) = \sum_{\kappa=0}^{k} \binom{k}{\kappa} D^\kappa f^i \cdot D^{k-\kappa} g^j.$$

If $x$ is a complex number, $\mu$ an integer, such that $f$, $g$ and their derivatives are defined at $\mu x$, we let

$$A_{i,j}^{k;\mu}(t) = \sum_{\kappa=0}^{k} \binom{k}{\kappa} D^\kappa f^i(\mu t) \cdot D^{k-\kappa} g^j(\mu x).$$

We shall assume from now on that $f(t) = t$, and that $x$ is a period of $g$ and of its derivatives. We let $g_{k,j} = D^k g^j(0)$. Then

$$A_{i,j}^{k,\mu}(t) = \sum_{\kappa=0}^{k} \binom{k}{\kappa} \mu^{i-\kappa} i(i-1)\cdots(i-\kappa+1)t^{i-\kappa} g_{k-\kappa,j}$$

is a polynomial in $t$, and

$$D^k F(\mu x) = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \varphi_{ij} A_{i,j}^{k,\mu}(x).$$

THEOREM 3. *Let $k$, $\mu$ be such that*

$$0 \leqq k \leqq n-1 \quad \textit{and} \quad 0 \leqq \mu \leqq m-1.$$

*Let $\Delta_g = \mathrm{Det}(g_{kj})$, and let*

$$\Delta_1 = \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2^{m-1} \\ \vdots & \vdots & & \vdots \\ 1 & m-1 & \cdots & (m-1)^{m-1} \end{vmatrix}$$

*be the Vandermonde determinant. Then*

$$\mathrm{Det}\,|A_{i,j}^{k,\mu}(t)| = t^{\frac{1}{2}mn(m-1)}\,\Delta_1^n\,\Delta_g^m.$$

*(Here, $(i, j)$ indexes rows, and $(k, \mu)$ indexes columns.)*

*Proof.* Since each $A_{i,j}^{k,\mu}$ is a polynomial of degree $\leqq i$, it follows that

$$\Delta(t) = \mathrm{Det}\,A_{i,j}^{k,\mu}(t)$$

is a polynomial of degree $\leqq \frac{1}{2}mn(m-1)$. We shall prove that

$$\Delta(t) = ct^{\frac{1}{2}mn(m-1)}$$

for some constant $c$, and determine this constant.

Let $H = (H^1, \ldots, H^M)$ be an $M \times M$ matrix of functions, with column vectors $H^1, \ldots, H^M$. Then for any integer $s \geqq 0$, we have

$$D^s\,\mathrm{Det}(H) = \sum_{(\sigma)} \mathrm{Det}(D^{\sigma_1}H^1, \ldots, D^{\sigma_M}H^M),$$

where the sum is taken over all $(\sigma)$ such that $\sigma_1 + \cdots + \sigma_M = s$. This is trivially proved by induction.

We contend that $D^s\Delta(0) = 0$ if $s < \frac{1}{2}mn(m-1)$. *Proof:* Let $M = mn$. For any $s$ with $0 \leqq s \leqq \frac{1}{2}mn(m-1)$, we have

$$D^s\Delta(0) = \sum_{(\sigma)} \operatorname{Det}(\ldots, D^{\sigma_{k\mu}}A^{k\mu}(0), \ldots) = \sum_{(\sigma)} \operatorname{Det}_{(\sigma)},$$

where the sum is taken over $(\sigma)$ such that $\sum\sigma_{k\mu} = s$. We have, for any integer $\sigma \geqq 0$,

$$D^\sigma A^{k\mu}_{ij}(0) = \begin{cases} \binom{k}{i-\sigma}\mu^\sigma i!g_{k-i+\sigma,j} & \text{if } \sigma \leqq i \\ 0 & \text{if } \sigma > i. \end{cases}$$

If a term $\operatorname{Det}_{(\sigma)}$ is such that, for some $k$, there exist $\mu_1 \neq \mu_2$ such that $\sigma_{k\mu_1} = \sigma_{k\mu_2} = \tau$, then we can factor out $\mu_1^\tau$ and $\mu_2$ from the $(k, \mu_1)$ and $(k, \mu_2)$ columns respectively, and obtain a determinant in which two distinct columns are equal. Thus $\operatorname{Det}_{(\sigma)} = 0$. Hence the only non-zero determinants $\operatorname{Det}_{(\sigma)}$ are such that $\sigma_{k\mu_1} \neq \sigma_{k\mu_2}$ if $\mu_1 \neq \mu_2$. For such $(\sigma)$,

$$\sum_{\mu=0}^{m-1} \sigma_{k\mu} \geqq \frac{m(m-1)}{2}.$$

Taking the sum over $k$, we find that if $D^s\Delta(0) \neq 0$, then $s \geqq \frac{1}{2}mn(m-1)$. This proves the first part of our theorem, that $\Delta(t) = ct^{\frac{1}{2}mn(m-1)}$.

We must now determine the constant $c$. For this, we need a formal lemma on determinants.

Lemma. *Let*

$$X = (x_{kj}), \qquad k, j = 0, \ldots, n-1$$

*and*

$$Y = (y_{i\mu}), \qquad i, \mu = 0, \ldots, m-1$$

*be two matrices of independent variables, and let*

$$X * Y = (x_{kj}y_{i\mu})$$

*be the corresponding $mn \times mn$ matrix. Then*

$$\operatorname{Det}(X * Y) = \operatorname{Det}(X)^m \operatorname{Det}(Y)^n.$$

*Proof.* We may write

$$\operatorname{Det}(X * Y) = \operatorname{Det}\begin{vmatrix} x_{00}Y & x_{01}Y & \cdots & x_{0,n-1}Y \\ x_{10}Y & x_{11}Y & \cdots & x_{1,n-1}Y \\ \vdots & \vdots & & \vdots \\ x_{n-1,0}Y & x_{n-1,1}Y & \cdots & x_{n-1,n-1}Y \end{vmatrix}$$

Subtracting a multiple of the first column from each other column, we find that the determinant of $X * Y$ is equal to the determinant

$$\mathrm{Det}\begin{vmatrix} x_{00}Y & 0 & \cdots & 0 \\ x_{10}Y & x'_{11}Y & \cdots & x'_{1,n-1}Y \\ \vdots & \vdots & & \vdots \\ x_{n-1,0}Y & x'_{n-1,1}Y & \cdots & x'_{n-1,n-1}Y \end{vmatrix}$$

and by induction, we find

$$\mathrm{Det}(X * Y) = G(X)\,\mathrm{Det}(Y)^n,$$

where $G(X)$ does not depend on $Y$. By symmetry, it follows that

$$\mathrm{Det}(X * Y) = c_0\,\mathrm{Det}(X)^m\,\mathrm{Det}(Y)^n$$

for some constant $c_0$. Letting $X, Y$ be the unit matrices, we see that $c_0 = 1$. This proves our lemma.

Theorem 3 now follows at once from the lemma, if we simply take the determinant of the matrix obtained from $A_{ij}^{k\mu}$ by replacing each polynomial by its term of highest degree. It is then clear that the determinant $c$ is of the type considered in the lemma.

COROLLARY. *Let $f(t) = t$ and $g(t) = e^t$. Let $x = 2\pi\sqrt{-1}$. Then*

$$\mathrm{Det}\,|D^k(f^ig^j)(\mu x)| \neq 0.$$

## §6. *A transcendence measure for logarithms*

We shall illustrate a general method of Feldman by a special case, to see how accurate a result one can obtain under the most special hypotheses.

THEOREM 4. *For all $d$, $h \geqq 3$ and all algebraic numbers $\xi$ of degree $\leqq d$ and absolute height $\leqq h$, we have*

$$\log|2\pi i - \xi| \gg -d(d\log d + h)\log(d\log d + h).$$

*(The constant implicit in the symbol $\gg$ is an absolute constant.)*

*Proof.* Let $c$ be a constant, taken sufficiently large with respect to the absolute constant $C_1$ of Lemma 7, and with respect to $2\pi$. Let $\lambda$ be a parameter, of which we need only that it is sufficiently large with respect to $c$. Finally, let

$$N = \max\left(d, \frac{h}{\log h}\right)\cdot$$

We may assume $N$ sufficiently large with respect to $\lambda$.

Now let
$$B = e^{c\lambda^4 dN(\log N)^2}.$$

For convenience, let $x = 2\pi\sqrt{-1}$, and suppose that

$$\log|x - \xi| \leqq -\lambda^6\, dN(\log N)^2. \tag{1}$$

We shall reach a contradiction. Let $d_0$ be the precise degree of $\xi$. We consider a function in the usual manner,

$$F(t) = \sum_{i_0=0}^{d_0-1} \sum_{i=0}^{r_1-1} \sum_{j=0}^{r_2-1} a_{(i)} \xi^{i_0} t^i e^{jt}$$

with integer coefficients $a_{(i)} = a_{i_0 ij}$. We take

$$r_1 = [\lambda^3 d \log N] \qquad \text{and} \qquad r_2 = [\lambda^2 N \log N].$$

We want

$$|D^k F(\nu x)| < \frac{1}{B} \tag{2}$$

for

$$0 \leqq k \leqq [\lambda^3 N \log N] \qquad \text{and} \qquad 0 \leqq \nu \leqq [\lambda d \log N].$$

These conditions amount to solving linear inequalities, with

$$\text{number of variables} \gg dr_1r_2 \gg \lambda^5\, d^2N(\log N)^2,$$
$$\text{number of inequalities} \gg \lambda^4\, dN(\log N)^2.$$

Furthermore, easy estimates of the usual type show that

$$\log|\text{coefficients}| \ll \lambda^3 N(\log N)^2.$$

Hence by Lemma 5 on linear inequalities, there exists a solution in integers $a_{(i)}$ not all zero such that

$$\log|a_{(i)}| \ll \frac{1}{\lambda d} \log B \ll \lambda^3 N(\log N)^2.$$

We may consider $D^k F(\nu x)$ as a polynomial in two variables, say

$$D^k F(\nu x) = P_{k,\nu}(\xi, x).$$

Estimates of the usual type yield:

$$\log|P_{k,\nu}| \ll \lambda^3 N(\log N)^2,$$
$$\deg P_{k,\nu} \text{ in } \xi \ll d,$$
$$\deg P_{k,\nu} \text{ in } x \ll r_1 \ll \lambda^3 d \log N.$$

By Theorem 1, if $P_{k,\nu}(\xi, \xi) \neq 0$, then

$$-\lambda^3\, dN(\log N)^2 \ll \log |P_{k,\nu}(\xi, \xi)|. \tag{3}$$

On the other hand,

$$P_{k,\nu}(\xi, \xi) = P_{k,\nu}(\xi, x) + \int_x^\xi P'_{k,\nu}(\xi, z)\, dz. \tag{4}$$

Hence

$$|P_{k,\nu}(\xi, \xi)| \leqq |P_{k,\nu}(\xi, x)| + |x - \xi| M_{k,\nu}$$

where $M_{k,\nu}$ is a bound for the expression inside the integral sign, estimated by the same type of bound that we obtained above for the coefficients of our linear inequalities, namely

$$\log M_{k,\nu} \ll \lambda^3 N(\log N)^2.$$

Using (1) and (2), we find

$$\log |P_{k,\nu}(\xi, \xi)| \ll -\lambda^4\, dN(\log N)^2,$$

which contradicts (3) as soon as $\lambda$ is sufficiently large (with respect to an absolute constant). Hence

$$P_{k,\nu}(\xi, \xi) = 0$$

for all $k$, $\nu$.

We now wish to prove that

$$P_{r,\mu}(\xi, \xi) = 0$$

for

$$0 \leqq r \leqq r_2 - 1 \qquad \text{and} \qquad 0 \leqq \mu \leqq r_1 - 1.$$

If we can do this, then with the notation of the preceding section, we have

$$0 = P_{r,\mu}(\xi, \xi) = \sum_{i=0}^{r_1-1} \sum_{j=0}^{r_2-1} \varphi_{ij}(\xi) A_{ij}^{r\mu}(\xi).$$

By the Corollary of Theorem 3, the determinant of this system of linear equations is not 0, whence $\varphi_{ij}(\xi) = 0$ for all $i, j$, and we obtain a contradiction. We carry out this program in two steps.

*First step.* Let $\tau$ be a positive number such that $3 + \tau < 4$ but $3 + \tau + 1 + \tau > 5$. For instance $\tau = 3/4$. We shall prove that

$$P_{k',\nu'}(\xi, \xi) = 0$$

for

$$0 \leqq k' \leqq [\lambda^{3+\tau} N \log N] \qquad \text{and} \qquad 0 \leqq \nu' \leqq [\lambda^{1+\tau} d \log N].$$

We apply the interpolation Lemma 7, with $z' = \nu' x$, $r = k'$, $z_1, \ldots, z_m$ equal to the numbers $\nu x$ with $0 \leqq \nu \leqq [\lambda d \log N]$, omitting $z'$, and

$$R_1 = 4\lambda^{1+\tau} d(\log N)|x|,$$
$$R = 3C_1 R_1.$$

Since $r \log r \ll \lambda^{3+\tau} N(\log N)^2$, and

$$\log |F|_R \ll \lambda^{3+\tau}\, dN(\log N)^2,$$

the first term in the estimate of Lemma 7 is dominated by $(C_1 R_1/R)^{ml}$, whose logarithm is $\ll -\lambda^4\, dN(\log N)^2$. The second term satisfies a similar estimate, because of our choice of $B$, with a large constant $c$, so that in the product

$$\frac{2^r r!}{R_1^r} C_1^{ml} ml(1/\delta)^{ml} \max_{k,\nu} |D^k F(\nu x)|$$

the term involving derivatives dominates the estimate. Thus finally,

$$\log |D^{k'} F(\nu' x)| \ll -\lambda^4\, dN(\log N)^2.$$

Now we use again the estimate

$$|P_{k',\nu'}(\xi, \xi)| \leqq |P_{k',\nu'}(\xi, x)| + |x - \xi| M_{k',\nu'}$$

and argue as before. If $P_{k',\nu'}(\xi, \xi) \neq 0$, we get the inequalities

$$-\lambda^{3+\tau}\, dN(\log N)^2 \ll \log |P_{k',\nu'}(\xi, \xi)| \ll -\lambda^4\, dN(\log N)^2,$$

a contradiction which shows that $P_{k',\nu'}(\xi, \xi) = 0$.

We now obtain from (4) (with $k'$, $\nu'$ instead of $k$, $\nu$):

$$|P_{k',\nu'}(\xi, x)| \leqq |x - \xi| M_{k',\nu'},$$

whence

$$\log |D^{k'} F(\nu' x)| \ll -\lambda^6\, dN(\log N)^2. \tag{5}$$

*Second step.* We now prove that

$$P_{r,\mu}(\xi, \xi) = 0$$

for

$$0 \leqq r \leqq r_2 - 1 \qquad \text{and} \qquad 0 \leqq \mu \leqq r_1 - 1.$$

We use Lemma 7 again to estimate $D^r F(\mu x)$. We take $z' = \mu x$. We let

$z_1, \ldots, z_m$ be the numbers $\nu' x$, omitting $z'$, and

$$R_1 = 4r_1|x| = 4[\lambda^3 d \log N]|x|,$$
$$R = 3C_1 R_1.$$

Then $\log |F|_R \ll \lambda^5 \, dN(\log N)^2$ and hence by our choice of $\tau$, the first term is dominated by $(C_1 R_1/R)^{ml}$, whose logarithm is

$$\ll -\lambda^{4+2\tau} \, dN(\log N)^2.$$

For the second term, we use (5), and find that the logarithm of the second term is dominated by (5). Hence

$$\log |D^\tau F(\mu x)| \ll -\lambda^{4+2\tau} \, dN(\log N)^2$$

whence

$$\log |P_{\tau,\mu}(\xi, \xi)| \ll -\lambda^{4+2\tau} \, dN(\log N)^2.$$

On the other hand, estimating the degree and height of $P_{\tau,\mu}$, if

$$P_{\tau,\mu}(\xi, \xi) \neq 0,$$

then

$$-\lambda^3 \, dN(\log N)^2 \ll \log |P_{\tau,\mu}(\xi, \xi)|,$$

which is a contradiction. This proves the theorem.

Using the same method as that for Theorem 1, Feldman obtains:

THEOREM 5. *Let $\alpha$ be algebraic, and $x = \log \alpha \neq 0$. For all $d$, $h \geqq 3$, and all algebraic numbers $\xi$ of degree $\leqq d$ and absolute height $\leqq h$, we have*

$$\log |x - \xi| \gg - d^2(d \log d + h)(\log d) \log(d \log d + h).$$

The main difference is the appearance of $d^2$, due to the fact that one parameter is lost since $e^x = \alpha$, and this introduces an extra term, so that instead of having a polynomial $P_{k,\nu}$ in two variables, we have a polynomial $P_{k,\nu}(\xi, x, \alpha)$ in three variables. It is then necessary to adjust the values for $i$, $j$, $k$, $\nu$ accordingly. Also, the symbol $\gg$ now depends on the given $\alpha$.

Feldman also obtains analogous results for the Weierstrass $\wp$-function. In that case, the fact that such a function is of order 2 simply introduces another parameter, and makes the final result (using the same method) correspondingly worse. It should be noted, however, that Feldman proves the corresponding non-vanishing of the determinant of §5. It is an interesting problem, independent of the theory of transcendental numbers, to investigate such determinants for abelian functions and other generalized exponential functions. Finally, it should be mentioned that when

$h$ is large compared to $d$, then Feldman improves the dependence of the inequality on $h$, and obtains the following typical result.

THEOREM 6. *Let $\alpha$ be algebraic $\neq 0$, and $x = \log \alpha \neq 0$. For all $d \geqq 3$, and for all algebraic numbers $\xi$ of degree $\leqq d$ and absolute height $\leqq h$, with $h > d^4$, we have*

$$\log |x - \xi| \gg -h\, d^2(\log d)^2,$$

*where $\gg$ depends only on $\alpha$.*

## Historical note

This entire chapter is due to Feldman who obtained the results, and related ones, in the series of papers listed in the bibliography. These call for a number of comments.

We shall proceed systematically, and first make some very general comments on transcendental numbers and diophantine approximations.

The theory of transcendental numbers determines which classical numbers are linearly independent or algebraically independent (over the rationals). This requires a definition of the notion of classical number, and essentially, a classical number is one which appears as a value of a classical function suitably normalized. Let us give examples. In a classical situation, one meets an open subset of some complex space, say $U$, and a map $f: U \to V$ of $U$ into an algebraic variety. Given such a map, we can generate a field of numbers, by taking the smallest field $\Omega$ generated from the rational numbers by performing the following operations inductively, and iterating them:

Taking algebraic closure.

Adjoining values of $f$ and its inverse function with the argument in the field obtained inductively after a finite number of steps.

Examples of maps $f$ are given by exponential maps, uniformizing maps, solutions of algebraic differential equations, zeta functions (including $L$-series and gamma functions), etc.

Given say real numbers $x_1, \ldots, x_m$ in a field $\Omega$, one wishes to study the inequality

$$(*) \qquad |q_0 + q_1x_1 + \cdots + q_mx_m| < \frac{1}{q^m g(q)} \qquad (q = \max |q_i|)$$

with integers $q_i$, and some function $g$, positive and increasing. Assuming that $1, x_1, \ldots, x_m$ are linearly independent over the rationals, one defines $x_1, \ldots, x_m$ to be of *type* $\leqq g$ if the above inequality has only a finite number of solutions. Similarly, if $x$ is a given transcendental number,

one should say that it has transcendence type $\leqq g$ if the inequality (*) has only a finite number of solutions uniformly for $x_i = x^i$, and all $m$. Similarly, we can make a definition with respect to approximation by an algebraic number of degree $d$, using the inequality

$$(**) \qquad |x - \xi| < \frac{1}{H(\xi)^{d+1}g(H(\xi),d(\xi))} \qquad \text{with } H(\xi) = e^{h(\xi)}.$$

The results of [20] show that the theory of diophantine approximations of a number, say by rational numbers, achieves coherence and simplicity only if one measures the order of approximation not in the exponent but as a factor of $H^{d+1}$. (It is not entirely clear conjecturally to what extent $g$ should be independent of $d$.)

The general problem is now to determine inductively the type of a classical number obtained as value of classical function. One must of course first determine a type for algebraic numbers, and in this respect, the Thue-Siegel-Roth theorem appears as rather weak, in spite of the difficulties which one has encountered historically in reaching a proof for it.

Even in the case of algebraic numbers, no result is known to take into account varying degrees, say a result of type

$$|\alpha - \xi| \gg \frac{1}{H(\xi)^{d+1+\epsilon}}$$

or some such exponent as $d + 1 + \epsilon$, let alone more refined results. Such a result is not even known if the degree $d$ of $\xi$ is kept *fixed*.

The Feldman result may be seen as a first step in the inductive procedure, and the Liouville estimate (Theorem 1) is the induction hypothesis. However, given the present structure of the proof, this Feldman estimate is weaker than what should be expected. Thus it is good in that the exponent for $h$ and $d$ (in Theorem 4) is precisely equal to 1, but bad in that an extra log appears, so that one does not even get an estimate like

$$\log |x - \xi| \gg dh.$$

This, however, would only be a first step towards determining the refined types as in (*) and (**) above.

Even assuming a very good type for a number $x$, and using the more complicated techniques of Feldman, applied to the inductive procedure of Chapter V, I still do not see how to achieve a best possible result which would for instance yield the algebraic independence of $e$ and $\pi$, assuming that $e$ or $\pi$ has a very good transcendence type. It is very hard to tell whether this is because of a superficial defect in the proof, or whether one needs an entirely different structure for the algebraic independence proof.

It should be noted that Feldman's results (and a subsequent one by Gelfond [13], following Feldman's method) are the only ones which exhibit a good dependence on the degree $d$. For instance, the dependence on $d$ in Mahler's papers [24], using similar techniques to Siegel's, give a much worse dependence on $d$. We shall mention this again in the next chapter. The dependence on $h$ is much better in all known cases.

It may be that to achieve the best possible dependence on $d$, one must first determine a best possible type for the approximation by rational numbers, and *then* use a best possible type for approximation of algebraic numbers by rational numbers. Thus the induction must start with loaded hypotheses, not only going from numbers to values of a function, but also from the rational numbers to algebraic numbers. The few examples which one has now do suggest an absolutely fantastic rigidity in the entire theory.

## CHAPTER VII

# Linear Differential Equations

This chapter will deal with a method of Siegel, which is particularly efficient when the functions under consideration, aside from satisfying a (linear) differential equation, have a power series expansion of a special type. We shall begin by describing this type.

### §1. *E-functions*

An *E-function* is a function which admits a power series expansion

$$f(z) = \sum_{n=0}^{\infty} \alpha_n \frac{z^n}{n!}$$

with complex coefficients $\alpha_n$, belonging to some number field $K$, satisfying the following conditions:

E 1. *We have* $\|\alpha_n\| \leqq c^n$ *for some constant* $c$.

E 2. *There exists a sequence of integers* $d_n \in \mathbf{Z}$, $d_n > 0$ *such that* $d_n$ *is a denominator for* $\alpha_k$ $(k = 0, \ldots, n)$ *and*

$$d_n \leqq c^n.$$

(*Note:* We take the bound $c^n$ for convenience. Actually, everything goes through if we define $E$-functions using the bound $O(n^{\epsilon n})$ for every $\epsilon > 0$. However, all the examples satisfy the stronger conditions stated above, and I see no point here in introducing an extra parameter.)

The ordinary exponential function $e^z$ is an $E$-function. So is the Bessel function

$$J_0(z) = \sum \frac{z^{2n}}{(n!)^2}$$

or the Bessel function $J_\lambda$ with algebraic parameter $\lambda$. Similarly, hypergeometric functions (with algebraic parameter) are also examples of $E$-functions. We refer the reader to Siegel's book for such examples. Of course, polynomials with algebraic coefficients are $E$-functions.

In what follows, we assume that all $E$-functions mentioned have coefficients in the number field $K$. If $f$ is an $E$-function as above, we define

$$\mathrm{size}_n(f) = \mathrm{size}(\alpha_0, \ldots, \alpha_n)$$

to be the size of its first $n + 1$ coefficients (i.e. coefficients of $z^n/n!$). If $f'$ is the derivative of $f$, then

$$\mathrm{size}_n(f') \leqq \mathrm{size}_{n+1}(f),$$

whence in particular, $f'$ is also an $E$-function.

Let

$$g(z) = \sum \beta_n \frac{z^n}{n!}$$

be an $E$-function. Then $f + g$ is an $E$-function, and

$$\mathrm{size}_n(f + g) \leqq \mathrm{size}_n(f) + \mathrm{size}_n(g) + 2.$$

Furthermore,

$$f(z)g(z) = \sum \gamma_n \frac{z^n}{n!}$$

where

$$\gamma_n = \sum_{k=0}^{n} \binom{n}{k} \alpha_k \beta_{n-k},$$

whence $fg$ is an $E$-function, with

$$\mathrm{size}_n(fg) \leqq \mathrm{size}_n(f) + \mathrm{size}_n(g) + 2n,$$

because if $d_n$ is a denominator for $\alpha_0, \ldots, \alpha_n$ and $d_n'$ is a denominator for $\beta_0, \ldots, \beta_n$ then $d_n d_n'$ is a denominator for $\gamma_0, \ldots, \gamma_n$.

Finally, if $\alpha$ is in $K$ and $f$ is as above, an $E$-function, then $f(\alpha z)$ is also an $E$-function, with

$$\mathrm{size}_n f(\alpha z) \leqq \mathrm{size}_n(f) + n \cdot \mathrm{size}(\alpha).$$

## §2. *The Lindemann theorem*

We shall carry out a special case of the Siegel method to prove the Lindemann theorem.

THEOREM 1. *Let $\alpha_1, \ldots, \alpha_s$ be algebraic numbers, linearly independent over the rationals. Then $e^{\alpha_1}, \ldots, e^{\alpha_s}$ are algebraically independent.*

The proof will use several lemmas.

Let $K = \mathbf{Q}(\alpha_1, \ldots, \alpha_s)$. Let $\beta_1, \ldots, \beta_m$ be distinct non-zero elements of $K$, and let $E_j$ $(j = 1, \ldots, m)$ be the $m$ functions

$$E_j(z) = e^{\beta_j z}.$$

We shall form a new function

$$F_1(z) = P_1(z)e^{\beta_1 z} + \cdots + P_m(z)e^{\beta_m z}$$

with polynomials $P_1, \ldots, P_m$ having coefficients in $I_K$, not all zero, such that $F_1$ has a high zero at 0. We regard $m$ as given, and we shall deal with a parameter $n$. Constants $c, c_1, c_2, \ldots$ thus depend on the $\beta_j$ and $m$.

LEMMA 1. *Given an integer $n > 0$ we can find polynomials $P_j \in I_K[z]$ not all zero, such that:*

(i) *$\deg P_j < 2n$ and $\|P_j\| \leqq c^n n^{2n}$.*

(ii) *The function $F_1$ has a zero of order $\geqq (2m - 1)n$ at 0.*

(iii) *If*

$$F_1(z) = \sum_{\nu=0}^{\infty} a_\nu \frac{z^\nu}{\nu!}$$

*then $|a_\nu| \leqq c^\nu c^n n^{2n}$.*

*Proof.* We write $P_j$ with unknown coefficients, namely

$$P_j(z) = (2n-1)! \sum_{\mu=0}^{2n-1} x_{j\mu} \frac{z^\mu}{\mu!}$$

and have

$$E_j(z) = \sum_{\mu=0}^{\infty} \beta_j^\nu \frac{z^\nu}{\nu!} = \sum_{\nu=0}^{\infty} \beta_{j\nu} \frac{z^\nu}{\nu!}.$$

Then

$$P_j E_j(z) = (2n-1)! \sum_{\nu=0}^{\infty} b_{j\nu} \frac{z^\nu}{\nu!}$$

where

$$b_{j\nu} = \sum_{k=0}^{2n-1} \binom{\nu}{k} x_{jk} \beta_{j,\nu-k}.$$

We obtain

$$P_1E_1 + \cdots + P_mE_m = \sum_{\nu=0}^{\infty} a_\nu(x) \frac{z^\nu}{\nu!}$$

where

$$a_\nu(x) = (2n-1)!(b_{1\nu} + \cdots + b_{m\nu}).$$

We must solve for $(x)$ the linear equations $a_\nu(x) = 0$, with $\nu < (2m-1)n$.

We have $2mn$ unknowns

$$x_{j\mu} \quad \text{with } j = 1, \dots, m \text{ and } \mu = 0, \dots, 2n - 1$$

and $(2m - 1)n$ equations. The coefficients of these equations are in $K$, bounded by

$$\begin{aligned} \|\text{coefficients}\| &\leqq m \cdot \max \binom{\nu}{k} \max \|\beta_{j,\nu-k}\| \\ &\leqq c_1^n, \end{aligned}$$

since the binomial coefficient is bounded by $2^{(2m-1)n}$. A denominator for these coefficients is also bounded by $c_1^n$. By Siegel's lemma on linear equations, we can find a non-trivial solution with $x_{j\mu} \in I_K$ satisfying $\|x_{j\mu}\| \leqq c^n$. It is now an easy matter to estimate the coefficients $a_\nu$ for all $\nu$ to get the estimate (iii). Of course, only those $a_\nu$ may be $\neq 0$ for $\nu \geqq (2m - 1)n$. This proves our lemma.

Lemma 2. *Let $E_j$, $P_j$, $F_1$ be as in Lemma 1. Let*

$$F_{k+1} = D^k F_1$$

*where $D$ is the derivative, $k = 1, 2, \dots$ Write*

$$F_k = P_{k1}E_1 + \cdots + P_{km}E_m$$

*with polynomials $P_{kj}$. Then the rank of the matrix $(P_{kj})$ $(k, j = 1, \dots, m)$ is equal to $m$.*

*Proof.* Let $Y$ be the vertical vector of functions $(E_1, \dots, E_m)$. Then $Y$ satisfies the linear differential equation

$$Y' = QY$$

where $Q$ is the matrix

$$Q = \begin{pmatrix} \beta_1 & \cdots & 0 \\ \vdots & \beta_2 & & \vdots \\ & & \ddots & \\ 0 & \cdots & & \beta_m \end{pmatrix}$$

In fact, we have

$$D^k F_1 = (D + \beta_1)^k P_1 E_1 + \cdots + (D + \beta_m)^k P_m E_m.$$

Thus the matrix $(P_{kj})$ is none other than

$$\begin{pmatrix} P_1 & \cdots & P_m \\ (D + \beta_1)P_1 & \cdots & (D + \beta_m)P_m \\ \vdots & \cdots & \vdots \\ (D + \beta_1)^{m-1}P_1 & \cdots & (D + \beta_m)^{m-1}P_m \end{pmatrix}$$

Let $\Delta = \Delta(z)$ be its determinant. Then $\Delta$ is a polynomial in $z$, and we shall prove that it is not zero by looking at the highest power of $z$ occurring in it. In fact, let $u_1, \ldots, u_m$ be the leading coefficients of the polynomials $P_1, \ldots, P_m$ respectively, and let $d_1, \ldots, d_m$ be their degrees. Then our determinant has one term of type

$$\begin{vmatrix} u_1 & \cdots & u_m \\ \beta_1 u_1 & \cdots & \beta_m u_m \\ \vdots & \cdots & \vdots \\ \beta_1^{m-1} u_1 & \cdots & \beta_m^{m-1} u_m \end{vmatrix} z^{d_1 + \cdots + d_m}$$

plus other terms in the expansion which have lower degree. We factor out $u_1 \cdots u_m$ from the determinant, and see that the remaining constant is a Vandermonde determinant which is not 0. This proves Lemma 2.

As in the proof of Lemma 2, let $\Delta$ be the determinant

$$\Delta = \det(P_{kj}) \qquad (k, j = 1, \ldots, m).$$

Then

$$\deg \Delta \leqq (2n - 1)m.$$

Let $P$ be the matrix $(P_{kj})$ $(k, j = 1, \ldots, m)$, and let $\tilde{P}$ be the matrix such that

$$\tilde{P}P = \Delta I.$$

Thus $\tilde{P}$ is the transpose of the matrix of minors of $\Delta$. Let $F$ be the column vector of $(F_1, \ldots, F_m)$ and let $Y$ be as before, the column vector of $(E_1, \ldots, E_m)$. Then

$$F = PY \qquad \text{and} \qquad \Delta Y = \tilde{P}F.$$

Each $F_j$ $(j = 1, \ldots, m)$ has a zero at 0 of order $\geqq (2m - 1)n - m$, and since none of the components of $Y$ vanishes at 0, we conclude that

$$\operatorname{ord} \Delta \geqq (2m - 1)n - m.$$

(Here, ord means order at 0.) Comparing this order with the degree of $\Delta$, it follows that if $\xi$ is any complex number $\neq 0$, then

$$\begin{aligned} \operatorname{ord}_\xi \Delta &\leqq \deg \Delta - \operatorname{ord} \Delta \\ &\leqq n. \end{aligned}$$

LEMMA 3. *For any complex number $\xi \neq 0$, the matrix*

$$(P_{kj}(\xi)) \qquad (k = 1, \ldots, m + n \text{ and } j = 1, \ldots, m)$$

*has rank $m$.*

*Proof.* We have $\tilde{P}P = \Delta I$. For any $k = 1, 2, \ldots$ we have

$$(D + Q)^k P_{(1)} = P_{(k)},$$

where $P_{(k)}$ is the $k$-th row of $P$ viewed as column vector. Let $r = \operatorname{ord}_\xi \Delta$. We apply $(D + Q)^r$ to $\Delta I$, and find

$$(D + Q)^r(\Delta I) = \sum_{\mu=0}^{r} C_\mu (D + Q)^{\mu\, t}P,$$

where $C_\mu$ are matrices of polynomials. Evaluating these expressions at $\xi$, we see that in the expansion on the left, all terms will vanish except $D^r\Delta(\xi)I$, whence

$$D^r\, \Delta(\xi) I = \sum_{\mu=0}^{r} C_\mu(\xi)(D + Q)^{\mu\, t}P(\xi).$$

On the left we have a non-zero scalar matrix. On the right, the columns of the matrix $(D + Q)^{\mu\, t}P(\xi)$ are simply the columns

$$P_{(k)}(\xi) = {}^t(P_{k1}(\xi), \ldots, P_{km}(\xi))$$

with $k \leqq m + n$. It follows that these columns have rank at least $m$, thereby proving our lemma.

Let $\alpha \in K$, $\alpha \neq 0$. We shall estimate $|F_k(\alpha)|$ and $\|P_{kj}(\alpha)\|$. Further constants depend on the size of $\alpha$.

LEMMA 4. *Let $k \leqq m + n$. Assume $n \geqq c_2(\alpha)$. Then*

$$|F_k(\alpha)| \leqq c_3^n n^{3n} n^{-(2m-2)n}$$

$$\|P_{kj}(\alpha)\| \leqq c_3^n n^{3n}, \qquad \textit{and} \qquad \operatorname{den}(P_{kj}(\alpha)) \leqq c_3^n.$$

*Proof.* The first inequality will come from the estimates of the coefficients of $F_1$, applying $k$ derivatives, and using the fact that $D^kF_1$ begins with a high power of $z$, and hence a high factorial in the denominator which contributes the term $n^{-(2m-2)n}$ tending to 0 with $n$. We do this in detail.

The power series for $F_1$ is dominated term by term by

$$F_1(z) \prec c^n n^{2n} \sum_{\nu=(2m-1)n}^{\infty} c^\nu \frac{z^\nu}{\nu!}$$

and that of $F_k$ is therefore dominated by

$$F_k(z) \prec c^n n^{2n} c^k \sum_{\nu=(2m-1)n}^{\infty} c^{\nu-k} \frac{z^{\nu-k}}{(\nu - k)!}.$$

Therefore,

$$|F_k(\alpha)| \leqq c_4^n n^{2n} \sum_{\nu=(2m-1)n}^{\infty} \frac{(c|\alpha|)^{\nu-k}}{(\nu-k)!}.$$

We observe that for any integer $r > 0$,

$$\sum_{\nu=r}^{\infty} \frac{w^\nu}{\nu!} = \frac{w^r}{r!}\left(1 + \frac{w}{r+1} + \frac{w^2}{(r+1)(r+2)} + \cdots\right).$$

Here we take $r = (2m-1)n - k$. In making our estimate, we take the maximum value of $r$ when estimating numerators, and its minimum value when estimating denominators, for $0 \leqq k \leqq m + n$. We have

$$(2m-2)(n-1) \leqq (2m-2)n - m \leqq r \leqq (2m-1)n.$$

Since we took $n$ large compared to $\alpha$, the sum in parentheses is $\leqq 2$. Also, $r!$ is approximately equal to $r^r e^{-r}$. Putting all this together, we obtain the desired estimate for $|F_k(\alpha)|$. In fact, we get an exponent $n^{2n}$ instead of $n^{3n}$, but we have nevertheless put $n^{3n}$ to fit some later generalization.

To estimate $\|P_{kj}(\alpha)\|$, we estimate the size of the coefficients of $P_{kj}$. We know that

$$P_{(k)} = (D + Q)^k P_{(1)}.$$

It is easy to estimate this by induction, using the same technique as in Chapter III, §2, except that the situation here is easier. A given polynomial $P_j$ $(j = 1, \ldots, m)$ is dominated by

$$P_j(z) \prec c^n n^{2n}(1 + z)^{2n-1},$$

and applying $(D + \beta_j)^k$ to this polynomial is easily seen by induction to be dominated by

$$(D + \beta_j)^k P_j \prec c_5^n n^{3n}(1 + z)^{2n-1}.$$

(The extra power of $n$ comes from a term bounded by $(2n-1)^{m+n}$ arising from successive derivatives.) Substituting $\alpha$ for $z$ then gives the desired result. The estimate for the denominator of $P_{kj}(\alpha)$ is even more trivial.

The next, and final, lemma is the decisive step in the proof. From the linear independence of the *functions* $E_1, \ldots, E_m$ over the polynomials, it gives a lower bound for the rank of the *numbers* $E_1(\alpha), \ldots, E_m(\alpha)$ over $K$.

LEMMA 5. *The rank of $E_1(\alpha), \ldots, E_m(\alpha)$ over $K$ is $\geqq m/2[K : \mathbf{Q}]$.*

*Proof.* Let $r$ be this rank. Let

$$\begin{aligned} 0 &= \lambda_{11}E_1(\alpha) + \cdots + \lambda_{1m}E_m(\alpha) \\ &\vdots \\ 0 &= \lambda_{m-r,1}E_1(\alpha) + \cdots + \lambda_{m-r,m}E_m(\alpha) \end{aligned}$$

be $m - r$ linearly independent relations with coefficients $\lambda_{kj} \in I_K$. By Lemma 3, we can find $r$ functions among the $F_k$, $k \leqq m + n$, such that, if we put

$$\begin{aligned} F_{k_1}(\alpha) &= P_{k_1 1}(\alpha)E_1(\alpha) + \cdots + P_{k_1 m}(\alpha)E_m(\alpha) \\ &\vdots \\ F_{k_r}(\alpha) &= P_{k_r 1}(\alpha)E_1(\alpha) + \cdots + P_{k_r m}(\alpha)E_m(\alpha) \end{aligned}$$

then the matrix consisting of the $(\lambda)$ and the $(P_{k_ij}(\alpha))$ has rank $m$. Let $\delta$ be the determinant of this matrix. Then $\delta$ is an element of $K$, $\delta \neq 0$. We obtain

$$\delta E_1(\alpha) = B_1 F_{k_1}(\alpha) + \cdots + B_r F_{k_r}(\alpha)$$

where $B_1, \ldots, B_r$ are minors of the determinant $\delta$. From Lemma 4, we then have

$$\operatorname{size}(\delta) \leqq 3nr \log n + O(n) \qquad \text{and} \qquad \operatorname{den}(\delta) \leqq O(n).$$

Again by Lemma 4, we obtain the upper bounds

$$|B_1|, \ldots, |B_r| \leqq c_6^n n^{3n(r-1)}.$$

Lemma 4 also gives us an upper bound for $|F_k(\alpha)|$. We therefore obtain the upper bound

$$\begin{aligned} \log|\delta| &\leqq 3n(r-1)\log n + 3n\log n - (2m-2)n\log n + O(n). \\ &\leqq 3rn\log n - (2m-2)n\log n + O(n). \end{aligned}$$

We compare this with the size, divide by $n \log n$ throughout, get rid of $O(n)$, and find

$$2m - 2 \leqq 3r\,[K : \mathbf{Q}].$$

The assertion of the lemma follows trivially.

To prove Theorem 1, we start with the $s$ functions

$$f_1(t) = e^{\alpha_1 t}, \ldots, f_s(t) = e^{\alpha_s t}.$$

Let $g$ be a polynomial with coefficients in $K$, not all zero, of degree $d$. We must show that for $\alpha = 1$,

$$g(e^{\alpha_1}, \ldots, e^{\alpha_s}) = g\big(f_1(\alpha), \ldots, f_s(\alpha)\big) \neq 0.$$

Let $\nu$ be a large integer, and let

$$m_\nu = \binom{\nu + s}{s}$$

be the binomial coefficient.

There are precisely $m_\nu$ monomials

$$f_1^{\nu_1} \cdots f_s^{\nu_s}, \qquad {}_1 + \cdots + \nu_s \leqq \nu.$$

We let $m = m_\nu$ and let $E_1, \ldots, E_m$ be these monomials. (Thus, in our special case of the exponential function, the numbers $\beta_1, \ldots, \beta_m$ are the linear combinations

$$\nu_1\alpha_1 + \cdots + \nu_s\alpha_s, \qquad \nu_1 + \cdots + \nu_s \leqq \nu.)$$

If $g(f_1(\alpha), \ldots, f_s(\alpha)) = 0$, then for $\nu_1 + \cdots + \nu_s \leqq \nu - d$, we have

$$f_1(\alpha)^{\nu_1} \cdots f_s(\alpha)^{\nu_s} g(f_1(\alpha), \ldots, f_s(\alpha)) = 0.$$

In this way, we obtain relations with coefficients in $K$ among the $m_\nu$ monomials

$$f_1(\alpha)^{\nu_1} \cdots f_s(\alpha)^{\nu_s}, \qquad \nu_1 + \cdots + \nu_s \leqq \nu.$$

It is immediately seen that these relations are linearly independent over $K$, and we have

$$m_{\nu-d} = \binom{\nu + s - d}{s}$$

such relations. By Lemma 5, we must have

$$m_\nu - m_{\nu-d} \geqq m_\nu/2\,[K : \mathbf{Q}].$$

This is impossible, because $m_\nu$ and $m_{\nu-d}$ are both polynomials in $\nu$, starting with the same term $\nu^m/m!$. This contradicts the assumption that $g(f_1(\alpha), \ldots, f_s(\alpha)) = 0$, thereby proving Theorem 1.

## §3. *Shidlovsky's lemma*

In order to extend the proof of Theorem 1 to arbitrary $E$-functions satisfying a linear differential equation with rational functions as coefficients, we must state and prove a lemma which allows us to generalize Lemma 2 above. This is the only difficult point in extending the proof of Theorem 1, but involves only linear algebra, no arithmetic. Thus in this section, we may assume that $K$ is an arbitrary field of characteristic 0, and we deal with power series in $K[[z]]$.

We begin by some remarks on linear differential equations. Let

$$Y = {}^t(y_1, \dots, y_m)$$

be a column vector of power series, and assume that $Y$ satisfies the linear differential equation

$$Y' = QY$$

where $Q$ is a matrix $(Q_{ij})$ of rational functions in $K(z)$. Let $T = T(z)$ be the polynomial which is the greatest common denominator of the $Q_{ij}$. We call $T$ a polynomial denominator for $Q$ (or for the $Q_{ij}$). If $P_1, \dots, P_m$ are polynomials in $K[z]$, we let $P_{(1)}$ be their column vector. Let

$$F_1 = P_1 y_1 + \cdots + P_m y_m = \langle P_{(1)}, y \rangle$$

be the scalar product of $P_{(1)}$ and $Y$. We construct $F_k$ inductively by taking

$$F_{k+1} = TDF_k = (TD)^k F_1.$$

Here, $D = d/dz$ is the formal derivative of power series with respect to $z$. We see trivially that

$$\begin{aligned} D\langle P_{(1)}, Y\rangle &= \langle DP_{(1)}, Y\rangle + \langle P_{(1)}, DY\rangle \\ &= \langle DP_{(1)}, Y\rangle + \langle P_{(1)}, QY\rangle \\ &= \langle DP_{(1)}, Y\rangle + \langle {}^tQP_{(1)}, Y\rangle \\ &= \langle D + {}^tQ)P_{(1)}, Y\rangle. \end{aligned}$$

Hence

$$TDF_1 = \langle T(D + {}^tQ)P_{(1)}, Y\rangle$$

and, inductively,

$$F_k = \langle (T(D + {}^tQ))^k P_{(1)}, Y\rangle.$$

Thus we can write

$$F_k = P_{k1}y_1 + \cdots + P_{km}y_m$$

with polynomials $P_{kj}$.

SHIDLOVSKY'S LEMMA. *Let $y_1, \dots, y_m$ be as above, formal power series linearly independent over $K(z)$, and satisfying the linear differential equation $Y' = QY$, where $Q$ is a matrix of rational functions. Let $P_1, \dots, P_m$ be polynomials in $K[z]$, and let*

$$F_1 = P_1 y_1 + \cdots + P_m y_m.$$

*Let $T$ be a polynomial denominator for $Q$, and define inductively*

$$F_k = TDF_{k-1} = P_{k1}y_1 + \cdots + P_{km}y_m.$$

*Let $r$ be the rank of the matrix $(P_{kj})$ $(k, j = 1, \ldots, m)$ and suppose $r < m$. Then*

$$\operatorname{ord} F_1 \leqq r(\max \deg P_j) + c_0,$$

*where $c_0$ is a positive number depending only on $y_1, \ldots, y_m, Q$ and not on the $P_j$.*

Although the proof of the lemma is rather long, we shall not use any part of it later. We use only the statement of Shidlovsky's lemma, and then only in the proof of Lemma 2, §4. Hence the reader may omit the rest of this section without impairing his understanding of the rest of the chapter.

The proof will involve lemmas, numbered as 2.1, 2.2, ...

LEMMA 2.1. *Let $\varphi_1, \ldots, \varphi_n$ be power series in $K[[z]]$, and let $d$ be a positive integer. There exists an integer $N$ (depending on $\varphi$ and $d$) such that, if $P_1, \ldots, P_n$ are polynomials in $K[z]$ of degrees $\leqq d$, then either*

$$F = P_1\varphi_1 + \cdots + P_n\varphi_n$$

*is equal to* 0, *or*

$$\operatorname{ord} F \leqq N.$$

*Proof.* Consider first the case when we take the $P_1, \ldots, P_n$ to be constants in $K$. Write the column vector of functions $\Phi$:

$$\begin{array}{ccccccc} \varphi_1 & = & a_{10} & + & a_{11}z & + & a_{12}z_2 + \cdots \\ \vdots & & \vdots & & \vdots & & \vdots \\ \varphi_n & = & a_{n0} & + & a_{n1}z & + & a_{n2}z_2 + \cdots \end{array}$$

Let $C = (c_1, \ldots, c_n)$ be a constant vector, and let $A^0, A^1, \ldots$ be the column vectors of the matrix of coefficients of $\varphi_1, \ldots, \varphi_n$. We note that

$$C \cdot \Phi = c_1\varphi_1 + \cdots + c_n\varphi_n = 0$$

if and only if $C \cdot A^k = 0$ for $k = 0, 1, 2, \ldots$

Let $A^0, A^1, \ldots, A^M$ generate the space of column vectors. If $C \cdot \Phi \neq 0$, then at least one of the dot products $C \cdot A^k$ is not 0 for $0 \leqq k \leqq M$. This means that

$$\operatorname{ord} C \cdot \Phi \leqq M,$$

and proves our lemma in case the polynomials $P_1, \ldots, P_m$ are constant. The general case is reduced to this one by replacing $\varphi_1, \ldots, \varphi_m$ by $z^i\varphi_j$ $(i = 0, \ldots, d$ and $j = 1, \ldots, m)$.

LEMMA 2.2. *Let $V, W$ be two vector spaces consisting of power series, finite dimensional over the constant field $K$. There exists an integer $N$*

*such that if $\varphi, \psi$ are non-zero elements of $V, W$ respectively such that $\varphi/\psi$ is a rational function, then* $\deg \varphi/\psi \leqq N$.

(By the *degree* of a rational function, we mean the maximum of the degree of its numerator and denominator.)

*Proof.* Let $\{\varphi_1, \ldots, \varphi_n\}$ be a basis of $V$ over $K$, and let $\{\psi_1, \ldots, \psi_r\}$ be a basis of $W$ over $K$. Consider the set of constant vectors

$$C = (c_1, \ldots, c_n)$$

for which there exists a constant vector $C' = (c'_1, \ldots, c'_r)$ and a rational function $R$ such that

$$C \cdot \Phi = RC' \cdot \Psi.$$

Let $C_1, \ldots, C_s$ be a maximal set of linearly independent vectors $C$ over $K$. We can write

$$C_k \cdot \Phi = R_k C'_k \cdot \Psi \qquad (k = 1, \ldots, s)$$

with rational functions $R_k$. Given any $C$ in our set, written as

$$C = x_1 C_1 + \cdots + x_s C_s$$

with coefficients $x_i \in K$, we see that

$$\begin{aligned} C \cdot \Phi &= (x_1 R_1 C'_1 + \cdots + x_s R_s C'_s) \cdot \Psi \\ &= RC' \cdot \Psi. \end{aligned}$$

From this it is clear that the degree of $R$ is bounded in terms of the degrees of $R_1, \ldots, R_s$.

LEMMA 2.3. *Let $\varphi_1, \ldots, \varphi_m$ be power series, linearly independent over the constants. Then the Wronskian determinant*

$$W(\varphi_1, \ldots, \varphi_m) = \begin{vmatrix} \varphi_1 & \varphi_2 & \cdots & \varphi_m \\ \varphi'_1 & \varphi'_2 & \cdots & \varphi'_m \\ \vdots & \vdots & & \vdots \\ \varphi_1^{(m-1)} & \varphi_2^{(m-1)} & \cdots & \varphi_m^{(m-1)} \end{vmatrix}$$

*is not* 0.

*Proof.* This is a standard easy lemma on derivatives, which is proved by induction. We leave it to the reader.

COROLLARY. *Let $m > r$, and let $\varphi_1, \ldots, \varphi_m$ be solutions in $K[[z]]$ of the differential equation*

$$\psi_r D^r y + \psi_{r-1} D^{r-1} y + \cdots + \psi_1 y = 0,$$

*where* $\psi_1, \ldots, \psi_r \in K[[z]]$ *and* $\psi_r \neq 0$. *Then* $\varphi_1, \ldots, \varphi_m$ *are linearly dependent over the constants.*

We return to the differential equation, determined by the matrix $Q$. Let $P_1, \ldots, P_m$ be polynomials whose column vector is denoted by $P_{(1)}$. We form inductively

$$P_{(k+1)} = T(D + {}^tQ)P_{(k)}$$

and obtain the matrix $(P_{(1)}, \ldots, P_{(m)})$ whose transpose is written

$$\begin{pmatrix} P_{11} & \cdots & P_{1m} \\ \vdots & & \vdots \\ P_{m1} & \cdots & P_{mm} \end{pmatrix}.$$

Suppose that for some integer $r < m$ the vectors $P_{(1)}, \ldots, P_{(r)}$ are linearly independent, but $P_{(r+1)}$ can be written as a linear combination

$$P_{(r+1)} = g_1P_{(1)} + \cdots + g_rP_{(r)}$$

with rational functions $g_1, \ldots, g_r$. Applying $T(D + {}^tQ)$ to this expression, we find that $P_{(r+2)}$ can also be written as a linear combination of $P_{(1)}, \ldots, P_{(r)}$ with rational functions as coefficients. Inductively, it follows that the rank of the matrix $(P_{kj})$ $(k, j = 1, \ldots, m)$ is equal to $r$, and that its first $r$ rows are linearly independent.

LEMMA 2.4. *Given the matrix $Q$, there exists an integer $N$ having the following property. Let $P_1, \ldots, P_m$ be polynomials such that the rank of the matrix $(P_{kj})$ is $r < m$. After renumbering the indices, if necessary, suppose that*

$$\begin{pmatrix} P_{11} & \cdots & P_{1m} \\ \vdots & & \vdots \\ P_{r1} & \cdots & P_{rm} \end{pmatrix} = \begin{pmatrix} P_{11} & \cdots & P_{1r} & P_{1,r+1} & \cdots & P_{1m} \\ \vdots & & \vdots & \vdots & & \vdots \\ P_{r1} & \cdots & P_{rr} & P_{r,r+1} & \cdots & P_{rm} \end{pmatrix}$$

*is written in terms of two blocks, say*

$$(P_{\mathrm{I}}, P_{\mathrm{II}}),$$

*such that $P_{\mathrm{I}}$ is $r \times r$, $P_{\mathrm{II}}$ is $r \times (m - r)$, and $P_{\mathrm{I}}$ is non-singular. Then there is a matrix $A$ of rational functions such that*

$$P_{\mathrm{II}} = P_{\mathrm{I}}A$$

*and such that the degrees of the rational functions are $\leqq N$.*

*Proof.* Observe that this lemma is independent of any solutions of the differential equation. Consequently, after making a change in local param-

eter from $z$ to $z - z_0$, where $z_0$ is not a pole of the coefficients of $Q$, we may assume that we deal with solutions of the differential equation in the power series ring $K[[z]]$ and that 0 is not a pole of the coefficients of $Q$. Let $S_Q$ be the $K$-space of solutions of the differential equation $Y' = QY$ in vectors $Y$ with components in $K[[z]]$. Then $S_Q$ has dimension $m$ over the constant field $K$. In fact, each solution is determined uniquely by its initial condition (i.e. its value at 0), as is easily shown recursively, and the map $Y \mapsto Y(0)$ establishes the $K$-isomorphism between $S_Q$ and the space of $m$-tuples over $K$. Let $U = (U^{(1)}, \ldots, U^{(m)})$ be a fixed system of linearly independent column vectors of solutions, say those corresponding to the initial conditions

$$\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \cdots \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

We have, for any solution $Y$ of $Y' = QY$,

$$\begin{aligned} (TD)^r\langle P_{(1)}, Y\rangle &= \langle P_{(r+1)}, Y\rangle \\ &= g_1\langle P_{(1)}, Y\rangle + \cdots + g_r\langle P_{(r)}, Y\rangle \\ &= g_1\langle P_{(1)}, Y\rangle + \cdots + g_r(TD)^{r-1}\langle P_{(1)}, Y\rangle. \end{aligned}$$

This shows that $\langle P_{(1)}, Y\rangle$ satisfies a linear equation of order $r$, to which we can apply the corollary of Lemma 2.3. We conclude that the map

$$Y \mapsto \langle P_{(1)}, Y\rangle$$

is a $K$-linear map from the space $S_Q$ into a $K$-space of dimension $\leqq r$. Its kernel consists of those $Y$ which are orthogonal to $P_{(1)}$ (and hence to $P_{(k)}$ for all $k$). We can find a basis

$$\{\overline{Y}^{(1)}, \ldots, \overline{Y}^{(m)}\}$$

of $S_Q$ such that $\overline{Y}^{(1)}, \ldots, \overline{Y}^{(m-r)}$ consist of vectors orthogonal to

$$P_{(1)}, \ldots, P_{(r)}.$$

Hence we have

$$(P_{\mathrm{I}}, P_{\mathrm{II}})(\overline{Y}^{(1)}, \ldots, \overline{Y}^{(m-r)}) = 0.$$

There certainly is some matrix $A$ of rational functions such that

$$P_{\mathrm{II}} = P_{\mathrm{I}}A,$$

and we note that $A$ is $r \times (m - r)$, because the columns of $P_{\mathrm{II}}$ depend

on the columns of $P_{\mathrm{I}}$. We shall prove that $A$ is uniquely determined and can be expressed in a special way in terms of $\overline{Y}$, where $\overline{Y}$ is the matrix $(\overline{Y}^{(1)}, \ldots, \overline{Y}^{(m)})$.

Decompose $(\overline{Y}^{(1)}, \ldots, \overline{Y}^{(m-r)})$ into two vertical blocks,

$$(\overline{Y}^{(1)}, \ldots, \overline{Y}^{(m-r)}) = \begin{pmatrix} \overline{Y}_{\mathrm{I}}^{(1)} & \cdots & \overline{Y}_{\mathrm{I}}^{(m-r)} \\ \overline{Y}_{\mathrm{II}}^{(1)} & \cdots & \overline{Y}_{\mathrm{II}}^{(m-r)} \end{pmatrix},$$

such that the top block has $r$ rows and the bottom block has $m - r$ rows, corresponding to the decomposition of $(P_{\mathrm{I}}, P_{\mathrm{II}})$. Then we obtain

$$P_{\mathrm{I}}\overline{Y}_{\mathrm{I}}^{(k)} + P_{\mathrm{I}}A\overline{Y}_{\mathrm{II}}^{(k)} = 0 \qquad \text{for } k = 1, \ldots, m - r,$$

whence

$$(*) \qquad \overline{Y}_{\mathrm{I}}^{(k)} + A\overline{Y}_{\mathrm{II}}^{(k)} = 0 \qquad \text{for } k = 1, \ldots, m - r.$$

On the other hand, the columns

$$\overline{Y}_{\mathrm{II}}^{(1)}, \ldots, \overline{Y}_{\mathrm{II}}^{(m-r)}$$

are linearly independent, because the $m \times m$ matrix

$$\begin{pmatrix} \overline{Y}_{\mathrm{I}}^{(1)} & \cdots & \overline{Y}_{\mathrm{I}}^{(m-r)} & \cdots & \overline{Y}_{\mathrm{I}}^{(m)} \\ \overline{Y}_{\mathrm{II}}^{(1)} & \cdots & \overline{Y}_{\mathrm{II}}^{(m-r)} & \cdots & \overline{Y}_{\mathrm{II}}^{(m)} \end{pmatrix}$$

is non-singular, and a linear combination of the above-mentioned columns, together with relations (*), would yield a contradiction. Hence the matrix $A$ is uniquely determined. Relations (*) can then be viewed as a system of $r(m - r)$ linear equations for the components of $A$, having a unique solution.

There exists a constant $m \times m$ matrix $C$ such that

$$\overline{Y} = UC,$$

where $U$ is our fixed system of basic solutions of the original differential equation. If we substitute $UC$ for $\overline{Y}$ in the linear equations (*), we see that $A$ depends only on the polynomials $P_1, \ldots, P_m$ by means of the constant matrix $C$. In solving the system of linear equations for the components of $A$, we meet certain determinants involving the components of $\overline{Y}$. Let

$$\varphi_1, \ldots, \varphi_n$$

consist of all monomials of degree $\leqq m^2$ in the components of $U$. Then each component of $A$ can be expressed as a quotient of linear combinations with constant coefficients of $\varphi_1, \ldots, \varphi_n$. We can therefore apply Lemma 2.2 to conclude the proof of Lemma 2.4.

We shall now conclude the proof of Shidlovsky's lemma.

Let $(y_1, \ldots, y_m)$ be our solution of $Y' = QY$, linearly independent over $K(z)$, in the power series ring $K[[z]]$. Let

$$F_1 = P_1 y_1 + \cdots + P_m y_m$$

and

$$F_k = P_{k1} y_1 + \cdots + P_{km} y_m = (TD)^{k-1} F_1$$

as usual. Let $r$ be the rank of the matrix $(P_{kj})$ and assume $r < m$. Let $A = (A_{ij})$ be the matrix of Lemma 2.4, and let $T_1$ be a denominator for $A$, of bounded degree. By Lemma 2.4, we can write, for $k = 1, \ldots, r$ and $j = r + 1, \ldots, m$,

$$P_{kj} = \sum_{\nu=1}^{r} P_{k\nu} A_{\nu j}$$

and

$$F_k = P_{k1} y_1 + \cdots + P_{kr} y_r + \sum_{j=r+1}^{m} P_{kj} y_j,$$

whence

$$(**) \qquad F_k = \sum_{\nu=1}^{r} P_{k\nu} \left( y_\nu + \sum_{j=r+1}^{m} A_{\nu j} y_j \right).$$

Let $\Delta_0 = \det(P_{k\nu})$ $(k, \nu = 1, \ldots, r)$. Note trivially that by induction,

$$\deg P_{k\nu} \leqq \max \deg P_j + rq,$$

where $q$ is a constant depending only on the degrees of $Q_{ij}$. Consequently,

$$\operatorname{ord} \Delta_0 \leqq \deg \Delta_0 \leqq r \cdot \max \deg P_j + N_0$$

for some constant $N_0$. On the other hand, solving the linear equations (**) yields

$$\Delta_0 \left( y_\nu + \sum_{j=r+1}^{m} A_{\nu j} y_j \right) = \sum_{k=1}^{r} \Delta_{\nu k} F_k$$

where $\Delta_{\nu k}$ are subdeterminants of $\Delta_0$, and in any case are polynomials.

We multiply throughout by $T_1$ to clear denominators. We note that (from successive derivatives),

$$\operatorname{ord} F_1 - r \leqq \operatorname{ord} F_k,$$

and consequently

$$\operatorname{ord} F_1 - r \leqq \operatorname{ord} \Delta_0 + \operatorname{ord} \left( T_1 y + \sum_{j=r+1}^{m} T_1 A_{\nu j} y_j \right).$$

We can use Lemma 1 on the expression in parentheses on the right, to see that its order is bounded from above by a constant. Combining this with the upper bound for ord $\Delta_0$ in terms of its degree, we have finally

$$\operatorname{ord} F_1 \leqq r \cdot \max \deg P_j + N_1$$

for some constant $N_1$, thereby proving Shidlovsky's lemma.

## §4. *The general theorem*

We wish to extend Theorem 1 to arbitrary $E$-functions, satisfying a linear differential equation with rational functions as coefficients. Let $f_1, \dots, f_s$ be $E$-functions, satisfying the differential equation

$$X' = Q^*X,$$

where $X$ is a column vector of $(X_1, \dots, X_s)$, and

$$Q^* = (Q^*_{ij}) \qquad (i, j = 1, \dots, s)$$

is a square matrix of rational functions in $K(z)$ over the number field $K$. We shall prove the theorem of Siegel-Shidlovsky:

THEOREM 2. *Assume that $f_1, \dots, f_s$ are algebraically independent over $K(z)$, and satisfy a linear differential equation as above. Let $\alpha \in K$ be distinct from 0, and from the poles of the rational functions $Q^*_{ij}$. Then the values*

$$f_1(\alpha), \dots, f_s(\alpha)$$

*are algebraically independent.*

Examples of functions as in Theorem 2, besides the exponential function, are given by the Bessel functions, solutions of the differential equation

$$y'' + \frac{1}{z}y' + \left(1 - \frac{\lambda^2}{z^2}\right)y = 0,$$

the constant $\lambda$ being taken in a number field. If $J_\lambda$, $J'_\lambda$ are two linearly independent solutions, then it is known that they are algebraically independent (over $\mathbf{C}(z)$) if $2\lambda$ is not an odd integer. The proof of this fact can be found for instance in Siegel's book, and involves only function theory, no arithmetic. The proof is by no means dull, but the main point of the theory of transcendental numbers is to reduce the algebraic independence of *values* to the algebraic independence of *functions*, so that it is not unreasonable to omit such a function-theoretic proof in the present book, in view of its easy availability elsewhere. It is easy to verify that

$J_\lambda$, $J'_\lambda$ are $E$-functions, and thus we obtain a special case of Theorem 2, originally proved by Siegel, namely the algebraic independence of $J_\lambda(\alpha)$, $J'_\lambda(\alpha)$ whenever $\alpha$ is an algebraic number $\neq 0$, and $2\lambda$ is distinct from an odd integer.

We also refer to Siegel [32] for examples of hypergeometric functions satisfying the hypotheses of Theorem 2.

The proof of Theorem 2 follows very closely that of Theorem 1. There will be essentially no change in Lemma 1. For Lemma 2, we quote Shidlovsky's lemma, and then there is no further difficulty. Lemma 3 is essentially the same as before, and the estimates of Lemma 4 are only very slightly more difficult, since we must take polynomial denominators into account. We devote a small amount of space to these estimates. The final Lemma 5 does not change, and the last clinching arguments are exactly as before.

We shall repeat the lemmas in the general context, and preserve their numbering.

LEMMA 1. *Let $E_1, \ldots, E_m$ be E-functions, with coefficients in $K$, and let $n$ be a positive integer. There exist polynomials $P_1, \ldots, P_m \in I_K[z]$ not all zero, such that*

(i) $\deg P_j < 2n$ *and* $\|P_j\| \leqq c^n n^{2n}$.

(ii) *The function*

$$F_1 = P_1E_1 + \cdots + P_mE_m = \sum_{\nu=0}^{\infty} a_\nu \frac{z^\nu}{\nu!}$$

*has a zero of order* $\geqq (2m-1)n$ *at* 0.

(iii) *We have* $|a_\nu| \leqq c^\nu c^n n^{2n}$.

*Proof.* There is no change from the previous case.

Let $E_1, \ldots, E_m$ be $E$-functions with coefficients in $K$, and assume that they are linearly independent over $K(z)$. Assume also that they satisfy the linear differential equation

$$Y' = QY,$$

where $Q$ is a matrix $(Q_{ij})$ of rational functions in $K(z)$. Let $T$ be a polynomial denominator for $Q$, say $T \in I_K[z]$, and let

$$F_1 = P_1E_1 + \cdots + P_mE_m$$

be as in Lemma 1. We may then form inductively

$$F_k = TDF_{k-1} = P_{k1}E_1 + \cdots + P_{km}E_m$$

with polynomials $P_{kj}$.

LEMMA 2. *Let $E_j$, $P_j$, $F_1$ be as in Lemma 1, and assume that*

$$(E_1, \ldots, E_m)$$

*satisfy the linear differential equation*

$$Y' = QY$$

*as above. If $n > c_0'$ (where $c_0'$ depends only on $E_1, \ldots, E_m, Q$), then the rank of the matrix $(P_{kj})$ $(k, j = 1, \ldots, m)$ is equal to $m$.*

*Proof.* If the rank $r$ of $(P_{kj})$ is $< m$, we have, by Shidlovsky's lemma,

$$(2m - 1)n \leqq r(2n - 1) + c_0,$$

from which the assertion of Lemma 2 is obvious.

We let $\Delta$ be the determinant

$$\Delta = \det(P_{kj}) \qquad (k, j = 1, \ldots, m).$$

For some constant $c_1$ depending on $\deg T$ and $\deg TQ_{ij}$, we have

$$\deg \Delta \leqq (2n - 1)m + c_1.$$

Let $P$ be the matrix $(P_{kj})$ $(k, j = 1, \ldots, m)$ and let $\widetilde{P}$ be the matrix such that

$$\widetilde{P}P = \Delta I.$$

Let $F$ be the column vector of $(F_1, \ldots, F_m)$. Then

$$F = PY \qquad \text{and} \qquad \Delta Y = \widetilde{P}F.$$

Each $F_j$ $(j = 1, \ldots, m)$ has a zero at 0 of order $\geqq (2m - 1)n - m$, and $\operatorname{ord} E_j \leqq c_2$ for some constant $c_2$. Consequently

$$\operatorname{ord} \Delta \geqq (2m - 1)n - c_3.$$

Comparing this order with the degree of $\Delta$, it follows that if $\xi$ is any complex number $\neq 0$ then

$$\begin{aligned} \operatorname{ord}_\xi \Delta &\leqq \deg \Delta - \operatorname{ord} \Delta \\ &\leqq n + c_4. \end{aligned}$$

LEMMA 3. *For any complex number $\xi \neq 0$, and $\xi$ not equal to any zero of $T$, the matrix*

$$(P_{kj}(\xi)) \qquad (k = 1, \ldots, n + c_5, j = 1, \ldots, m)$$

*has rank $m$.*

*Proof.* For any $k$ we have

$$(T(D + {}^tQ))^k P_{(1)} = P_{(k)}.$$

Let $r = \operatorname{ord}_\xi \Delta$. We apply $(T(D + {}^tQ))^r$ to $\Delta I = {}^t(\tilde{P}P)$ and find as before an expression

$$(T(D + {}^tQ))^r(\Delta I) = \sum C_\mu (T(D + {}^tQ))^{\mu}\, {}^tP$$

where $C_\mu$ are matrices of polynomials. Evaluating these expressions at $\xi$, we have again

$$T^r(\xi) D^r \Delta(\xi) I = \sum_{\mu=0}^{r} C_\mu(\xi) (T(D + {}^tQ))^{\mu}\, {}^tP(\xi).$$

On the left we have a non-zero scalar matrix. On the right, the columns of the matrix $(T(D + {}^tQ))^{\mu}\, {}^tP(\xi)$ are simply the columns

$$P_{(k)}(\xi) = {}^t(P_{k1}(\xi), \ldots, P_{km}(\xi)),$$

with $k \leqq m + r$. It follows that these columns have rank at least $m$, thereby proving the lemma.

Let $\alpha$ be our element of $K$, distinct from 0 and the zeros of $T$. We shall estimate $|F_k(\alpha)|$ and $\|P_{kj}(\alpha)\|$.

LEMMA 4. *Let* $k \leqq n + c_5$. *Assume* $n \geqq c_6$. *Then*

$$|F_k(\alpha)| \leqq c_7^n n^{3n} n^{-(2m-2)n},$$
$$\|P_{kj}(\alpha)\| \leqq c_8^n n^{3n} \quad \text{and} \quad \operatorname{den} P_{kj}(\alpha) \leqq c_8^n.$$

*Proof.* First, we note that by induction, one proves easily that

$$(TD)^k = \sum_{\mu=0}^{k} T_\mu D^\mu$$

for all $k$, where $T_\mu$ is a polynomial dominated by

$$T_\mu(z) \prec c^k k!(1 + z)^{k+\deg T}.$$

The constant $c$ depends on $T$, of course. We apply $(TD)^k$ to $F_1$ for $k \leqq n + c_5$. We can easily estimate the derivative $D^\mu F_1(\alpha)$ as we did previously, to find an estimate of type

$$|D^\mu F_1(\alpha)| \leqq c^n n^{2n} n^{-(2m-2)n}.$$

The $k!$ contributes one more power $n^n$ giving a total of $n^{3n}$. The $|T_\mu(\alpha)|$ are trivially estimated, and we obtain the estimate for $|F_k(\alpha)|$ as stated

in the lemma. The estimates for the $P_{kj}(\alpha)$, which don't involve a power series, only polynomials, is even easier and is left to the reader.

LEMMA 5. *The rank of* $E_1(\alpha), \ldots, E_m(\alpha)$ *over* $K$ *is* $\geqq m/2\,[K : \mathbf{Q}]$.

*Proof.* The proof is identical with the proof in the special case. There is no need to repeat it.

To prove Theorem 2, we repeat essentially verbatim the final arguments of §2, with our given functions $f_1, \ldots, f_s$. We let again $E_1, \ldots, E_m$ be the monomials

$$f_1^{\nu_1} \cdots f_s^{\nu_s}, \qquad \nu_1 + \cdots + \nu_s \leqq \nu,$$

and let $\nu \to \infty$. If $g$ is a polynomial with coefficients in $K$, not all zero, of degree $d$, and if

$$g(f_1(\alpha), \ldots, f_s(\alpha)) = 0,$$

then we have

$$f_1^{\nu_1}(\alpha) \cdots f_s^{\nu_s}(\alpha) g(f_1(\alpha), \ldots, f_s(\alpha)) = 0$$

for $\nu_1 + \cdots + \nu_s \leqq \nu - d$. This is a system of linear relations, with coefficients in $K$, among the $m = m_\nu$ monomials

$$f_1^{\nu_1}(\alpha) \cdots f_s^{\nu_s}(\alpha), \qquad \nu_1 + \cdots + \nu_s \leqq \nu.$$

These relations are linearly independent over $K$, and we have $m_{\nu-d}$ such relations. Since $m_\nu - m_{\nu-d} \geqq m_\nu/2\,[K : \mathbf{Q}]$, we get the contradiction again by Lemma 5, thereby proving Theorem 2.

## §5. *A transcendence measure*

We shall see that by modifying the arguments at the end of the proof, one can obtain a result concerning a transcendence measure for the numbers $f_1(\alpha), \ldots, f_s(\alpha)$.

THEOREM 3. *Let* $f_1, \ldots, f_s$ *be E-functions, algebraically independent over* $K(z)$, *and satisfying a linear differential equation as at the beginning of §4. Let* $\alpha$ *be an algebraic number, distinct from* 0 *and from the poles of the rational functions* $Q_{ij}^*$. *Let* $g(X_1, \ldots, X_s)$ *be a polynomial with coefficients in* $\mathbf{Z}$, *of degree* $d$. *Then*

$$|g(f_1(\alpha), \ldots, f_s(\alpha))| \geqq c|g|^{-bd^s}$$

*where* $c$ *is a number* $> 0$, *depending on the* $f_j$, $s$, $Q^*$, $\alpha$, *and (unfortunately)* $d$, *while* $b$ *depends on the degree* $N = [K(\alpha) : \mathbf{Q}]$ *and on* $s$.

*Proof.* We start after Lemma 4. With ulterior motives, we select an integer $l > 0$ such that

$$1 > 4N\left(1 - \left[\frac{lN-1}{l}\right]^{\sigma}\right) \qquad \text{for all } \sigma,\ 1 \leqq \sigma \leqq s.$$

We let

$$m = \binom{s + lNd}{s} \quad \text{and} \quad v = \binom{s + lNd - d}{s}.$$

Let $w = m - v$. Then there are exactly $m$ monomials

$$f_1^{\nu_1} \cdots f_s^{\nu_s}$$

of degree $\leqq lNd$, and those monomials are $E$-functions, which we denote by $E_1, \ldots, E_m$. The differential equation $X' = Q^*X$ gives rise to a differential equation $Y' = QY$ with a matrix of rational functions $Q$.

There are exactly $v$ polynomials of type

$$f_1^{\nu_1} \cdots f_s^{\nu_s} g(f_1, \ldots, f_s),$$

with $\nu_1 + \cdots + \nu_s \leqq \lambda Nd - d$. We denote these functions by

$$\psi_1, \ldots, \psi_v.$$

We have inductively, using the construction of Lemma 1, with $n > c_0'$,

$$F_k = P_{k1}E_1 + \cdots + P_{km}E_m,$$

and Lemma 3 states that the matrix $(P_{kj}(\alpha))$ with $j = 1, \ldots, m$ and $k \leqq n + c_5$ has rank $m$ for some constant $c_5$. We can write

$$\begin{aligned} \psi_1 &= \lambda_{11}E_1 + \cdots + \lambda_{1m}E_m \\ &\vdots \\ \psi_v &= \lambda_{v1}E_1 + \cdots + \lambda_{vm}E_m, \end{aligned}$$

with coefficients $\lambda_{ij}$ which can be taken to be coefficients of the original polynomial $g$. Their absolute values are therefore bounded by $|g|$. We can find then $m - v = w$ functions among the $F_k$ ($k \leqq n + c_5$), say

$$\begin{aligned} \varphi_1 &= P_{k_1 1}(\alpha)E_1 + \cdots + P_{k_1 m}(\alpha)E_m \\ &\vdots \\ \varphi_w &= P_{k_w 1}(\alpha)E_1 + \cdots + P_{k_w m}(\alpha)E_m \end{aligned}$$

such that the determinant $\delta$ of the matrix consisting of the $(\lambda)$ and the $P_{k_ij}(\alpha)$ is not zero. We may assume without loss of generality that $E_1 = 1$,

namely the monomial of degree 0. Then

$$\delta = A_1\psi_1(\alpha) + \cdots + A_v\psi_v(\alpha) + B_1\varphi_1(\alpha) + \cdots + B_w\varphi_w(\alpha),$$

where the coefficients $(A)$ and $(B)$ are obvious subdeterminants of $\delta$. Each $P_{kj}$ has degree $\leqq 2n - 1 + (k-1)q$ where $q$ depends only on $Q$. It is easy to estimate the coefficients $(A)$ and $(B)$, and we find

$$|A_1\psi_1(\alpha) + \cdots + A_v\psi_v(\alpha)| \leqq c_9 g(f(\alpha))|g|^{v-1}n^{4nw},$$
$$|B_1\varphi_1(\alpha) + \cdots + B_w\varphi_w(\alpha)| \leqq c_{10}|g|^v n^{4nw} n^{-(2m-2)n}.$$

One sees easily that

$$\|\delta\| \leqq |g|^v n^{4nw},$$

and that a denominator for $\delta$ is bounded by $c^n$. By the usual lemma, we find

$$(*) \qquad 1 \leqq c_{11}|g|^{Nv} n^{4nwN}\left[\frac{|g(f(\alpha))|}{|g|} + \frac{1}{n^{n(2m-2)}}\right].$$

We take $n$ to be the smallest integer $> c_0'$, and such that

$$n^n > 2c_{11}|g|^N.$$

Recall that $w = m - v$. We contend that

$$2m - 2 - 4N(m - v) > v.$$

It will certainly suffice to prove that $m > 4N(m - v)$, in view of the inequality

$$2m - 2 - 4N(m - v) - v = m - 2 - (4N - 1)(m - v).$$

We have

$$s!m = (lNd + 1)\cdots(lNd + s) = (lN)^s d^s + \xi_{s-1}(lN)^{s-1} d^{s-1} + \cdots,$$
$$s!v = (lN - 1)^s d^s + \xi_{s-1}(lN - 1)^{s-1} d^{s-1} + \cdots,$$

where $\xi_{s-1}, \ldots, \xi_0$ are integers depending only on $s$. We estimate each

$$(lN)^\sigma - (lN - 1)^\sigma$$

using the definition of $l$, and our contention follows at once.

If we now substitute $2c_{11}|g|^N$ for $n^n$ we get an upper bound of $\frac{1}{2}$ for the second term on the right of $(*)$. Transposing this $\frac{1}{2}$ to the left of $(*)$ yields

$$|g(f(\alpha))| \geqq (|g|^{Nv} n^{4nwN})^{-1}$$

under our assumption that $n^n > 2c_{11}|g|^N$, or rewriting this in the log notation,

$$\log |g(f(\alpha))| \geqq -Nv \cdot \log |g| - 4nwN \cdot \log n.$$

Since $n$ is chosen smallest satisfying the inequality $n^n > 2c_{11}|g|^N$, we see that $n^n$ is of the order of magnitude of $|g|^N \log g$. Both $v$ and $w$ are obviously of type $b\, d^s$ with a suitable constant $b$, and hence

$$\log |g(f(\alpha))| \geqq -bN\, d^s \log |g| - bN^2\, d^s \log |g| - c_{12}.$$

This proves our theorem.

It is in fact easy to see that $w$ has degree $s - 1$ in both $N$ and $d$, and consequently that our constant $b$ is of type $b_0 N^{s+1}$ where $b_0$ depends only on $s$.

## *Historical note*

This entire chapter is due to Siegel (cf. [31] and [32]), except for Shidlovsky's lemma [30]. In his original paper, Siegel proved the non-vanishing of the crucial determinant only in special cases (including the case of the Bessel function), but left the general case open. He axiomatizes the situation in his book [32]. Shidlovsky saw how to prove the non-vanishing in general, and consequently closed the last gap in obtaining the general theorem stated here as Theorem 2.

In his original paper, Siegel also obtains a transcendence measure for the values of his functions, of the same type as Theorem 3. Once Shidlovsky proved his lemma, it was clear that Siegel's argument could be extended to the general case also (cf. [19]).

This estimate does not give a transcendence type as we defined it in Chapter V, because we view $d$ as fixed. I have checked that the proof of Theorem 1 in fact carries with it an explicit determination of the constant as a function of the degree, i.e. a function of $m$. In the estimates, these involve expressions like $m^{2m}$ or $m^{3m}$ (or a similar low exponent of $m^m$). This is of course very unfortunate. As far as I know, there is no result known on the approximation of $e^\alpha$ ($\alpha$ algebraic) by algebraic numbers $\xi$ which depend on $d$ in a manner similar to the Feldman results of Chapter VI.

It seems probable that to obtain good dependence on $d$, one will have first to improve (or rather change completely) the known methods used to prove approximation theorems concerning algebraic numbers, and apply similar methods to values of transcendental functions.

Finally, we note that Lindemann actually proves something slightly stronger than the algebraic independence of $e^{\alpha_1}, \ldots, e^{\alpha_s}$ if $\alpha_1, \ldots, \alpha_s$

are linearly independent over $\mathbf{Q}$. He proves that if $\beta_1, \ldots, \beta_m$ are distinct non-zero algebraic numbers, then $e^{\beta_1}, \ldots, e^{\beta_m}$ are linearly independent over the field of algebraic numbers. This does not come out of the Siegel method as it stands: The discrepancy is apparent in Lemma 5. Just to round out the theory, it would be worth while to see if one could not adjust the Siegel method so that Lemma 5 yields this stronger result, in the general case of $E$-functions satisfying a linear differential equation, in other words, if the functions are linearly independent, then their values are also linearly independent.

One also wishes to investigate transcendental numbers from the point of view of diophantine approximations. A general discussion is given in [21], and also in my book *Introduction to Diophantine Approximations.*

## APPENDIX

# The $p$-adic Case

The theory of transcendental numbers can also be developed over $p$-adic fields. We let $\mathbf{C}_p$ be the completion of the algebraic closure of the $p$-adic field $\mathbf{Q}_p$. Then $\mathbf{C}_p$ plays the role of the complex numbers.

In considering values of functions, the theorems are local, concerned with values of power series converging in some neighborhood of the origin. Extensions of theorems from the complex case to the $p$-adic case have proved themselves to be either fairly easily accessible in the past, with a proof which closely parallels the complex case, or completely out of reach. We shall now give examples of both cases.

The counting of zeros has an analogue in a theorem of Mahler [23]:

THEOREM. *Let* $f(t) = \sum a_\nu t^\nu$ *be a power series such that* $a_\nu \in \mathbf{C}_p$, $|a_\nu|_p \leqq 1$, *and* $\lim a_\nu = 0$. *Let* $m$ *be a positive integer, and*

$$x_1, \ldots, x_n \in \mathbf{C}_p$$

*such that* $|x_i|_p \leqq 1/p^m$. *Assume that* $f(x_i) = 0$ *for* $i = 1, \ldots, n$. *Then given* $x \in \mathbf{C}_p$ *with* $|x|_p \leqq 1/p^m$, *we have*

$$|f(x)|_p \leqq 1/p^{mn}.$$

*Furthermore, for the p-adic absolute value on the number field K, we have the estimate*

$$-\mathrm{size}(\alpha) \ll \log |\alpha|_p$$

*for all* $\alpha \neq 0$ *in* $K$.

The exponential function is defined by the usual series, and converges in the open disc of radius $p^{-1/(p-1)}$ in $\mathbf{C}_p$, where it satisfies the functional equation.

The results of Chapter I then go over without difficulty. We give Theorem 1 as an example.

THEOREM 1. *Let* $\beta_1, \beta_2 \in \mathbf{C}_p$ *be linearly independent over* $\mathbf{Q}$, *and* $z_\nu$ $(\nu = 1, 2, 3)$ *in* $\mathbf{C}_p$ *be also linearly independent over* $\mathbf{Q}$. *Assume that*

*$\beta_1$, $\beta_2$ have p-adic value $\leqq 1$ and that $|z_\nu|_p < 1/p$. Then at least one of the numbers*

$$\exp(\beta_1 z_\nu), \qquad \exp(\beta_2 z_\nu) \qquad (\nu = 1, 2, 3)$$

*is transcendental (over* $\mathbf{Q}$*).*

*Proof.* We take a large integer $c$ and let

$$f(t) = \exp(p^c \beta_1 t), \qquad g(t) = \exp(p^c \beta_2 t).$$

Then the power series for $f$, $g$ have $p$-integral coefficients, tending to 0, and it suffices to prove that not all values $f(z_\nu)$, $g(z_\nu)$ lie in $K$ (using the functional equation). Suppose the contrary. We form

$$F = \sum_{i,j=1}^{r} a_{ij} f^i g^j$$

with the same values of $r$ and $n$ as before, requiring $F$ to have zeros as before. Nothing is changed in this part of the proof which occurs entirely within the number field, and we obtain the same upper estimate for the coefficients $a_{ij}$, and the values of $F$. We let $s$ be as before, and now estimate $|F(w)|_p$. Applying the theorem of Mahler, we find

$$-s^{5/2} \ll \log |F(w)|_p \ll -s^3 \log p,$$

which gives the contradiction when $s$ is sufficiently large.

The theorems concerning algebraic groups go over in a similar manner. The exponential map can be represented locally by power series, with integral coefficients, tending to 0. We do not have any global conditions when dealing with abelian varieties. We state the theorem to give an example to the reader.

THEOREM 2. *Let $G$ be a linear group or an abelian variety defined over the field of algebraic numbers. Let $\varphi\colon D \to G_{\mathbf{C}_p}$ be a 1-parameter subgroup defined on a disc around the origin in $\mathbf{C}_p$. Let $\Gamma$ be a subgroup of $D$ having at least 3 linearly independent elements over $\mathbf{Q}$ in the linear case, and 5 such elements in the abelian case. If $\varphi(\Gamma)$ is contained in the group of algebraic points of $G$, then $\varphi$ parametrizes locally an algebraic subgroup of $G$ of dimension 1. In other words, there exists an open subdisc $D_0$ of $D$ containing 0, and an algebraic subgroup $H$ of $G$ of dimension 1 such that $\varphi(D_0)$ is an open subgroup of $H_{\mathbf{C}_p}$ containing the origin.*

In the situation of Theorem 2, one may say that the one parameter subgroup is locally algebraic. In the complex case, essentially by analytic continuation, the corresponding subgroup is algebraic.

In Chapter II, the situation is not quite as good. The results concerning $e^\alpha$ and $\alpha^\beta$ are still valid $p$-adically, and locally, wherever they make sense. Thus we must read $\exp(\alpha)$ instead of $e^\alpha$, and $\exp(\beta \log \alpha)$ instead of $\alpha^\beta$. Under the obvious conditions that $\beta$ and $\alpha$ lie in the proper domains for convergence of the exp and log functions, we have the same statements as in the complex case. This is due to Mahler [23]. The general theorem on differential equations is also true, as shown by Adams [1]. Here, one begins to feel some difficulties, since instead of a finite number of points $z_1, \ldots, z_m$, one must take an infinite sequence.

However, the analogue for the Weierstrass $\wp$-function is not known, and the difficulty here lies in the fact that it is of arithmetic order 2. In the complex case, we could take a large radius $R$ to estimate our function, but in the local $p$-adic case, this is impossible, and thus the question remains open.

An analogous difficulty arises for the transcendence of the Bessel function $J(\alpha)$ for algebraic $\alpha \neq 0$. The power series for the Bessel function converges just like the exponential series, but the technique of the factorials used by Siegel breaks down, and no other technique is known at present to replace it.

The $p$-adic theory of transcendental numbers has applications to various problems in algebraic number theory. We mention two of these. First, Leopoldt has defined a regulator in the $p$-adic case. Contrary to the complex case, no proof is known showing that his regulator is non-zero. This would follow from the algebraic independence of ($p$-adic) logarithms of multiplicatively independent algebraic numbers.

Second, Corollary 2 of Theorem 1, Chapter I, §1 answered a question put to me by Serre, which he encountered in his study of characters of idele classes, taking algebraic values. Of course, the same problem had occurred earlier (cf. the historical note of Chapter II, i.e. the paper of Alaoglu-Erdös).

# Bibliography


[1] W. Adams, "Transcendental numbers in the $p$-adic domain," *Am. J. Math.* (to appear).

[2] ——, "Asymptotic approximations to $e$," *Proc. Nat. Acad. Sciences USA* (1966), pp. 28–31.

[3] W. Baily, "On the theory of theta functions, the moduli of abelian varieties, and the moduli of curves," *Ann. of Math.* **75** (1962), pp. 342–381.

[4] N. I. Feldman, "The approximation of certain transcendental numbers, I," *Izvestia Akad. Nauk SSSR Ser. Math.* **15** (1951), pp. 53–74.

[5] ——, *idem*, II, *loc. cit.*, pp. 153–176.

[6] ——, "Arithmetic properties of values of elliptic functions," *Izvestia Akad. Nauk SSSR Ser. Math.* **22** (1958), pp. 563–576.

[7] ——, "Arithmetic properties of logarithms of algebraic numbers," *Izvestia Akad. Nauk SSSR Ser. Math.* **24** (1960), pp. 475–492.

[8] ——, "On the transcendental number $\pi$," *Izvestia Akad. Nauk SSSR* **24** (1960), pp. 357–368.

[9] ——, "Arithmetic properties of solutions of certain transcendental equations," *Vestnik*, Moscow, No. 1 (1964), pp. 13–20.

[10] —— and A. O. Gelfond, "On lower bounds of forms in three logarithms of algebraic numbers," *Vestnik*, Moscow, No. 5, (1949).

[11] —— and A. O. Gelfond, "On the measure of the relative transcendence of certain numbers," *Izvestia Akad. Nauk SSSR Ser. Math.* **14** (1950), pp. 493–500.

[12] A. O. Gelfond, "Sur les propriétés arithmétiques des fonctions entières," *Tohoku Math. J.* **30** (1929), pp. 280–285.

[13] ——, *Transcendental and algebraic numbers*, Dover, New York, 1960.

[14] C. Hermite, "Sur la fonction exponentielle," *Oeuvres III*, pp. 150–181.

[15] S. Lang, *Diophantine Geometry*, Interscience, New York, 1962.

[16] ——, "Transcendental points on group varieties," *Topology* **1** (1962), pp. 313–318.

[17] ——, "Algebraic values of meromorphic functions I," *Topology* **3** (1965), pp. 313–318, and II, to appear.

[18] ——, "Diophantine approximations on toruses," *Amer. J. of Math.* **86** (1964), pp. 521–533.

[19] ——, "A transcendence measure for $E$-functions," *Mathematika* **9** (1962), pp. 157–161.

[20] ——, "Asymptotic diophantine approximations," *Proc. Nat. Acad. Sciences USA* (1966), pp. 31–34.

[21] ——, "Report on diophantine approximations," *Bull. Soc. Math. France* **93** (1965), pp. 177–192.

[22] F. Lindemann, "Über die Zahl $\pi$," *Math. Annalen* **20** (1882), pp. 213–225.

[23] K. Mahler, "Über transzendente $p$-adische Zahlen," *Compositio Mathematica* **2** (1935), pp. 259–275.

[24] ——, "On the approximation of logarithms of algebraic numbers," *Philos. Trans. Roy. Soc. Londond Ser. A* **245** (1953), pp. 371–398.

[25] ——, "On the approximation of $\pi$," *Proceedings Koninklijke Nederl. Adak. Wetensch. Amsterdam. Ser. A* **56** (1953), pp. 30–42.

[26] A. Néron, "Quasi-fonctions et hauteur sur les variétés abeliennes," *Annals of Math.* (1965), pp. 349–381.

[27] T. Schneider, "Zur Theorie der Abelschen Funktionen und Integrale," *J. reine angew. Math.* (1941), pp. 110–128.

[28] ——, "Ein Satz über ganzwertige Funktionen als Prinzip fur Transzendenzbeweise," *Math. Ann.* **121** (1949–1950), pp. 131–140.

[29] ——, *Einführung in die transzendenten Zahlen*, Springer Verlag, Berlin, 1957.

[30] A. V. Shidlovsky, "On a criterion of algebraic independence," *Izvestia Akad. Nauk SSSR* **23** (1959), pp. 35–66.

[31] C. L. Siegel, "Über einige Anwendungen diophantischer Approximationen," *Abh. Preuss. Akad. Wiss.* (1929), pp. 1–41.

[32] ——, *Transcendental Numbers*, Annals of Math. Studies **16**, Princeton, 1949.

[33] A. Weil, *Variétés Kählériennes*, Hermann, Paris, 1958.

*Note:* The reader will find much more complete bibliographies in Gelfond's and Schneider's books. We have included here only the most significant papers.

ABCDE69876

## THE AUTHOR

Serge Lang received the Ph.D. degree from Princeton University and is currently Professor of Mathematics at Columbia University. Professor Lang is the author of *Algebra, Linear Algebra, Algebraic Numbers, A First Course in Calculus, A Second Course in Calculus, Algebraic Structures,* and *Introduction to Diophantine Approximations,* and the co-editor of *Collected Papers of Emil Artin* — all published by Addison-Wesley. His other books include *Abelian Varieties, Diophantine Geometry, Introduction to Algebraic Geometry,* and *Introduction to Differential Manifolds.*

PRINTED IN U.S.A.

## OTHER BOOKS OF INTEREST

### ALGEBRAIC STRUCTURES

BY SERGE LANG *In Press (1966)*

This book is designed for use as a text in modern algebra. Together with the author's *Linear Algebra,* it constitutes a curriculum for an algebra program addressed to undergraduates. It is, however, self-contained logically. Where overlap exists between this text and *Linear Algebra* — such as on groups of matrices and rings of endomorphisms — the emphasis in each text is quite different.

### LINEAR ALGEBRA

BY SERGE LANG *294 pp, 20 illus (1966)*

Starting with the basic notion of vectors in real Euclidean space, this book sets the pattern for much that follows. The book is intended as a text for a one-year course at the undergraduate level. Examples and special cases are worked out in abundance, and abstractions are developed slowly.

### ALGEBRA

BY SERGE LANG *508 pp (1965)*

This basic text for a one-year course in algebra at the graduate level assumes a familiarity with the basic language of mathematics (i.e. essentially sets and mappings), and the integers and rational numbers. Chapters are logically independent of each other, and there is maximum flexibility in the presentation of topics.

### ALGEBRAIC NUMBERS

BY SERGE LANG *163 pp, 2 illus (1964)*

The purpose of this book is to give an exposition of the classical and basic algebraic and analytic number theory. In a certain sense, the plan of the book is more or less that used by Hilbert in his *Bericht,* although both the algebraic and the analytic aspects of number theory have been updated.

ADDISON-WESLEY PUBLISHING COMPANY
READING, MASSACHUSETTS · PALO ALTO · LONDON · DON MILLS, ONTARIO